CENTRAL LIBRARY
CALCUTTA

Cuk-H05877-4-243679

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

4

370.1505
B777

FEBRUARY 1963

VOL. XXXIII, PART 1

Price One Pound

III

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
L. B. BIRCH

Associate Editors:
P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL

SIR FRED J. SCHONELL
C. W. VALENTINE
S. WISEMAN

Vol. XXXIII, Part I

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education, once a term, in February, June and November. The subscription is £2 a year post free payable in advance, or £1 each number.

Subscriptions and Orders should be sent to "The Manager, B.J.E.P. Department, Methuen and Co., Ltd., 36, Essex Street, Strand, London, W.C.2." Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to L. B. Birch, The Institute of Education, The University, Sheffield, 10. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to L. B. Birch, at the above address. *Advertisements* and *Subscriptions* for this *Journal* should be sent to Manager, B.J.E.P. Department, Methuen and Co., 36, Essex Street, Strand, London, W.C.2.

Published by
METHUEN AND CO., LTD., 36, ESSEX STREET, STRAND, LONDON, W.C.2.,
for the British Journal of Educational Psychology, Ltd.

Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of

The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

Edited by

L. B. BIRCH

Associate Editors:

P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
SIR FRED J. SCHONELL
C. W. VALENTINE
S. WISEMAN

Vol. XXXIII

1963

4

Published by
METHUEN AND CO., LTD., 36, ESSEX STREET, STRAND, LONDON, W.C.2.,
for the British Journal of Educational Psychology, Ltd.

Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of

The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

370.1505
B.777

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

EDITORIAL

THESIS ABSTRACTS

Since 1938, this *Journal* has regularly published abstracts of appropriate theses which have been accepted for higher qualifications in psychology and education. In recent years, however, in spite of the increased output of theses, these have come mainly from a very small number of universities. This note is included to remind Heads of Departments of Psychology and of Institutes and Departments of Education of the service which is available.

Thesis abstracts for consideration should be forwarded to the Editor by the Head of the Department concerned and may be intended either as the sole published account of the thesis or as a preliminary report of a research which is intended to be published at length at some future date. It is essential that they should be abstracts, not short articles, and Heads of Departments are specially asked to sub-edit them to make them as clear as possible consistent with brevity. Normally, abstracts of more than 1,500 words will not be considered and only exceptionally will tables be allowed.

Proofs will normally be returned to the Department of origin for correction.

CORRECTION

The heading of the three papers on researches using the Maudsley Personality Inventories in the last issue of the *Journal* (p. 241) should have been " Three Experiments using Maudsley Personality Inventories " not " Children's Personality Inventories." Savage used the M.P.I. in his research.

Dr. Warburton has also pointed out that the statement by Callard and Goodfellow that the J.M.P.I. is the only test of this kind available for English children, is not strictly true. Cattell's High School Personality Questionnaire (H.S.P.Q.) for ages 12-16 and his Children's Personality Questionnaire (C.P.Q.) for ages 8-12, are available in this country. Although they were standardised in the United States they have been used in Manchester and have been found suitable for use with English children.

VOLUME XXXIII FEBRUARY, 1963 PART I

A FURTHER STUDY OF THE EDUCATIONAL PROGRESS OF CHILDREN WHO HAD RECEIVED REMEDIAL EDUCATION

BY K. LOVELL, C. BYRNE AND B. RICHARDSON
(University of Leeds)

SUMMARY. Children who had received full-time teaching in Remedial Centres of an English education authority were traced, and their reading ages assessed along with their understanding of certain arithmetic concepts. A Summary of the data obtained is given together with the inter-relationships found between a number of the relevant variables

I.—INTRODUCTION.

IN an earlier paper (Lovell, *et al.* 1962) some details were given relating to the reading ages of children who had received remedial teaching and were subsequently followed-up. The pupils concerned in that study had received one period of remedial teaching per week, each period lasting 45 minutes. In the present study junior school children in a different education authority, who had been sent to full-time remedial centres for a period, were followed-up and their reading ages assessed. Furthermore, these children also undertook a test of arithmetic concepts at follow-up.

It must be clearly stated at the outset that this investigation throws no light on whether the children who attended these centres became better adjusted to school work and life, or whether their general motivation and morale improved under full-time remedial teaching. But the study does throw light on the likely educational progress and eventual performance, during school years, of children considered to be in need of remedial teaching. Furthermore, it is now possible to compare the progress of these children with those in the earlier study who had remedial teaching for only one period a week, and it has been possible to throw some further light on the question as to whether, in the long run, children who receive remedial teaching in special centres make more progress than those who stay in their own schools.

II.—SAMPLE STUDIED AND PROCEDURE.

The local education authority had five full-time remedial centres and the teachers-in-charge visited the local junior schools and picked out children, using intelligence and reading tests, who were considered to be in need of remedial teaching. Not all children so picked out attended the centres. Sometimes, the head of a junior school did not regard a particular child as suitable, and some parents refused permission for their children to attend the centres in lieu of school. In general, those who attended the centres were deemed to be the most likely to make most improvement under special teaching.

An examination of the records at the centres yielded complete data on 261 children who had attended full-time, and who had returned to ordinary schools not later than July, 1960. The follow-up study was done during the period November, 1961—March, 1962, so that no child was included in the survey unless he had been back at ordinary school for at least sixteen months.

The subjects were then divided up into those who were backward and those who were not, the former group being defined as those pupils whose reading ages (on whatever test was used at the time of referral) were less than 85 per cent. of their chronological ages. There was thus found to be 184 children classified as backward, and 77 as not backward. It will be quite obvious at this point that some children had been selected for attendance at these centres on the grounds of a discrepancy between reading age and mental age. Of the backward children, nine had moved to other parts of the country, one was in hospital, four were absent from school during the follow-up period, and six could not be traced by the authority, giving 164 backward children available for study. Of the non-backward group only one child was not found (she had moved to a private school outside the authority's area) giving 76 pupils in this category available for investigation.

The teachers-in-charge of the centres were free to use intelligence or reading tests (often the N.F.E.R. Sentence Reading Test) of their own choice in selecting children for the centres. Of the backward children fifty-nine had taken the Moray House Picture Test 1, fifty-four the Intelligence Tests for young children (Valentine), twenty-five the Sleight Non-Verbal Intelligence Test, fourteen the N.F.E.R. Non-Verbal Test 5, and twelve the Terman-Merrill revision of the Binet test. Again, either just before or immediately after admission, 142 subjects had taken the Schonell Graded Vocabulary Test R1 and twenty-two the Holborn Reading Scale for Infants and Juniors ; while at the time of return to school, 102 had been given the Schonell Graded Vocabulary Test R1, and sixty-two the Holborn Reading Scale.

In the non-backward group forty-eight had taken the Moray House Picture Test 1, eleven the Intelligence Tests for Young Children (Valentine), three the Sleight Non-Verbal test, thirteen the N.F.E.R. Non-Verbal Test 5, and one the Terman-Merrill revision of the Binet test. Just before or immediately after admission to the centres, sixty-two children had their reading ages assessed using the Schonell Graded Vocabulary Test R1 and fourteen the Holborn Reading Scale, while on return to school the numbers for these two tests were respectively fifty-seven and nineteen.

On follow-up, the children were found in over fifty schools. Both the Schonell Graded Vocabulary Test R1, and the Vernon Graded Word Reading Test were given to assess the reading ages. Moreover, each pupil undertook, individually, a test which involved the understanding of certain arithmetic concepts devised by one of the authors (K.L.). The questions, which were duplicated were presented to the subjects. The tester read aloud a question, and the child followed on the paper. The latter could either answer orally or in writing. No time limit was imposed. When the tester was satisfied that the child had answered the question to the best of his ability the experimenter passed on to the next question. The percentage numbers of children responding correctly to some of the least well answered questions are shown in Table 4.

In order to have a further check on educational progress, figures will be presented showing the numbers of children in the bottom stream, next from bottom stream, etc., of multi-stream schools, on follow-up (Table 5).

At one centre initial data were obtained on fifty children who, on the basis of a discrepancy between R.A. and M.A., had been regarded as being suitable for attendance at the centre, but who, for the kind of reasons already explained, did not, in fact, attend. Such children were left in ordinary schools and will be referred to as the control group. All children who were picked as likely to be suitable for this particular centre had taken the N.F.E.R. Sentence Reading Test ; in addition, other reading tests were given to children when they actually joined the centre. The control group was matched with pupils who went to the centre for chronological age, reading quotient (N.F.E.R. Sentence Reading Test), sex, and, as far as possible, intelligence test score and school ; and this matched group which attended the centre will be referred to as the experimental group. While the matching procedure was not perfect, it can be said with certainty that the teachers of the control group had no idea whatever, that some years later these children would be followed up and retested. Moreover, these fifty children in the experimental group were drawn from fifteen schools so that the classes out of which they were taken were but little reduced in size. At follow-up, then, the control group was also traced and the two reading tests given. See Table 6.

III.—RESULTS.

The data obtained have been given in tables for the sake of clarity.

TABLE 1

DATA RELATING TO THE PERIOD OF REMEDIAL TEACHING.

Variable	Backward Group N=164	Non-Backward Group N=76
I.Q.	M=98·8, σ =12·8	M=111·3, σ =13·3
R.A. (in years) at time of referral	M= 6·4	M= 8·2, σ = 1·0
C.A. (in years) at time of referral..	M= 9·2, σ = ·9	M= 8·7, σ = ·7
Mean Reading Quotient $\left(\frac{RA}{CA}\right)$	70	94
R.A. (in years) at end of remedial teaching	M= 8·4, σ =1·4	M=10·0, σ =1·3
C.A. (in years) at end of remedial teaching	M=10·4, σ =1·0	M= 9·6, σ = ·8
Mean Reading Quotient ..	81	104
Mean length of period of remedial teaching (years)	M=1·2	M= ·9
Extremes (in years) in period of remedial teaching	0·5–3·7	0·25–2·9
Mean improvement in R A. during remedial teaching (in years)....	2·0	1·8
Extremes (in years) of improvement during remedial teaching	Nil—5·4	·5–4·0

TABLE 2

DATA RELATING TO THE PERIOD SINCE REMEDIAL TEACHING CEASED.

Variable	Backward N=164	Non-Backward N=76
R.A. (in years) at follow-up (Schonell R1)	M=10·0, σ =1·8	11·4, σ =1·2
R.A (in years) at follow-up (Vernon Graded Word)	M= 9·8, σ =2·0	11·4, σ =1·5
C.A. (in years) at follow-up	M=13·1, σ =1·6	12·1, σ =1·3
Reading Quotient (R.A. from Schonell R1)	76	94
Mean progress (in years) since remedial teaching ceased	1·6	1·4
Mean length of time in years between end of remedial teaching and follow-up	2·7	2·5
Extremes of progress made (in years) since remedial teaching ended :		
12-23 months previously	−0·8 to 4·4 M=1·2 N=64	−0·5 to 2·4 M=0·7 N=28
24-35 months previously	−0·4 to 4·5 M=1·7 N=42	−0·6 to 4·1 M=1·8 N=30
36-47 months previously	0·5 to 3·2 M=2·0 N=31	1·0 to 3·8 M=2·2 N=12
48+ months previously	−1·8 to 5·8 M=1·9 N=27	0·4 to 2·3 M=1·1 N=6

TABLE 3

DISTRIBUTION OF READING AGES (SCHONELL R 1) AT FOLLOW-UP.

	Reading Age in Years										
	–5·4	5·5 — 6·4	6·5 — 7·4	7·5 — 8·4	8·5 — 9·4	9·5 — 10·4	10·5 — 11·4	11·5 — 12·4	12·5 — 13·4	13·5 — 14·4	
Backward	0	4	10	24	28	26	35	24	11	2	164
Non-Backward .. .	0	0	0	1	4	6	28	24	8	5	76

TABLE 4

PERCENTAGE NUMBERS OF CHILDREN GIVING CORRECT REPLIES TO SELECTED QUESTIONS IN THE TEST OF ARITHMETIC CONCEPTS.

Question	Backward	Non-Backward
1.—In the number 26 what does the figure 6 stand for (worth) ? Or, What does it tell us ?	37	68
2.—In the number 26 what does the figure 2 stand for ? Or, What does it tell us ?	44	66
3.—In the number 404, what does the middle figure 0 tell you ?	29	39
4.—In the number $\overset{*}{4}0\overset{\dagger}{4}$ how many times bigger is the 4 with a star than the 4 with a cross ?	10	13
5.—What number is half way between 26 and 32 ?..	39	61
6.—Write down the answer to this exercise without actually multiplying the numbers 9 × 57 divided by 9	13	22
6a.—Explain how you obtained the answer you have just given (in 6) Or, How did you get that ? ..	12	18

TABLE 5

PERCENTAGE NUMBERS OF CHILDREN IN GIVEN STREAMS OF MULTI-STREAM PRIMARY, SECONDARY MODERN OR ALL AGE SCHOOL AT FOLLOW-UP.

Group		Multi-Stream School			
	Single Stream School	Bottom Stream	Next to Bottom Stream in 3 or more Stream School	Top Stream	Intermediate Stream
Backward ..	16	59	18	5	2
Non-Backward	20	33	13	14	20

TABLE 6

DATA RELATING TO THE EXPERIMENTAL AND CONTROL GROUPS—50 IN EACH GROUP.

	Experimental Group	Control Group
I.Q.	M=116, σ =10·0	M=112, σ =10·6
C.A. (in years) when Experimental group first joined the Centre....	M=7·9	M=7·9
Reading Quotient (N.F.E.R. Sentence Reading)	M=85	M=85
Length of time (in years) at Centre	M=0·9	
R.A. (in years) at follow-up (Schonell R1)	M=10·7, σ =1·2	M=10·8, σ =1·3
R.A. (in years) at follow-up (Vernon)	M=10·5, σ =1·5	M=10·7, σ =1·5
C.A. (in years) at follow-up......	M=11·3	M=11·3
Reading Quotient (Schonell R1)..	94	96
Mean length of period (in years) since the experimental group first joined the Centre	M=3·4	M=3·4

TABLE 7

INTER-RELATIONSHIPS BETWEEN RELEVANT VARIABLES USING KENDALL'S TAU CO-EFFICIENT.

Relationship	Backward Group	Non-Backward Group
I.Q. and progress made during remedial teaching	·22**	·15*
I.Q. and R.A. (Schonell R.1) at follow-up	·27**	·10
Length of remedial teaching and progress made during remedial period	·21**	·35**
Progress made since remedial teaching ceased and time elapsed since	·27**	·30**
Progress made during remedial teaching and R.A. at follow-up	·37**	·24**
I.Q. and progress during remedial teaching holding constant the length of remedial teaching	·20†	·16†

*—Indicates significance at 5 per cent. level.

**—Indicates significance at 1 per cent. level.

†—It is not possible to test the significance of $\tau_{xy.z}$

IV.—DISCUSSION.

The results may be conveniently discussed under a number of headings :

(1) The data relating to the backward group (which has a mean I.Q. close to 100) during the period of remedial teaching shows much the same general trends as the corresponding data in the earlier study (Lovell *et al. op cit.*). There is a rise in the reading quotients of the same order of magnitude in both studies, and overall, we may say that as far as reading ages are concerned, the children in the earlier study, who had had remedial teaching for only one period a week, made equally as much progress per unit time as those now under discussion. Furthermore, the progress made in reading per unit time since remedial teaching ceased is almost identical in both investigations.

(2) There is tendency in the previous investigation for the reading quotients to fall between the end of treatment and follow-up. The same trend is clearly noticeable in both backward and non-backward groups in this study.

(3) Table 3 shows that only some 40 per cent. of the backward group had an understanding of place value involving tens and units at follow-up, and far fewer could fully appreciate place value in relation to hundreds, tens and units. This is a serious matter since an understanding of place value is fundamental in our number system.

(4) Over three-quarters of the children who were backward at time of joining the centres, were found at follow-up, in the bottom or next to bottom stream of a three or more stream school. A further 16 per cent. were in single stream schools, and only 5 per cent. in a top stream. This clearly confirms the view expressed in the earlier paper—that it is only very few children, backward at 8½–9 years of age, who are later found among the academically abler pupils. Not a single child of secondary school age in this group was found in a grammar school at follow-up.

(5) The findings given in Table 6 came somewhat as a surprise. After a mean period of almost 3½ years from the time of referral, no difference in the mean reading ages could be found between those who did, and those who did not attend the remedial centres. While the matching procedure could not be perfect, it was done with great care. Moreover, those who went to the centres were regarded as the most promising pupils. It is still possible, of course, that up to one year after referral, the mean reading age of those attending the centres might have been ahead of the mean reading age of those in school.

(6) The non-backward group (mean I.Q. 111) made about the same progress per unit time in reading as the backward group during the period of remedial teaching. The reading quotient rose but fell markedly again in the period after remedial teaching ended so that at follow-up the quotient was roughly the same as at the time of referral.

(7) In the arithmetic test some two-thirds of the non-backward group understood place value in relation to tens and units, but the more difficult questions relating to hundreds, tens and units remain poorly answered.

(8) In spite of the high mean I.Q. of the non-backward group, nearly half of them were found in the bottom stream, or next to bottom stream of a three or more stream school at follow-up. One in five of the group was found in a single stream school, and roughly one in seven in the top stream of a multi-stream school. But of those of secondary school age at follow-up, not a single child was found in a grammar school.

(9) It must be stressed that although the progress made by the majority of the children involved in this study was limited, the progress made was of great importance for the pupils concerned. To attain a reading age of 9+ years is a very real progress for the individual.

(10) The correlation coefficient between I.Q. and progress made during treatment (Table 7) is small but significant in the backward and non-backward group. This clearly parallels the finding in the earlier study.

(11) The previous study and this one, which together involved the follow-up of five hundred children, compared children taught singly, and in small groups, for one period per week, with children of somewhat higher I.Q. taught full-time in remedial centres. Whatever arrangements we have, the long-term results seem to be much the same. This surely suggests that we know very little as yet about the aetiology of scholastic failure and indicates the great need for research.

ACKNOWLEDGEMENTS.—Our thanks are due to the Chief Education Officer and the Senior Educational Psychologist of the Local Education Authority concerned for their generous help in conducting the enquiry and for permission to publish our findings.

V.—REFERENCES.

LOVELL, K., JOHNSON, E., PLATTS, D. (1962). A summary of a study of the reading ages of children who had been given remedial teaching. *Brit. J. Educ. Psychol*, 32, 66-71

(*Manuscript received, 10th May,* 1962.)

DEVELOPMENTAL PATTERNS OF SEVERELY SUBNORMAL CHILDREN

By MARY WOODWARD* AND DIANA J. STERN†

(*Fountain Hospital, London*)

SUMMARY. Eighty-three severely subnormal children aged under 9 years were observed with the aim of comparing their development in different areas and of examining the relation of various aspects of development to Piaget's sensori-motor stages. It was found that locomotor development was significantly in advance of speech development Children classified at various sensori-motor stages differed significantly in speech and in locomotor development. Certain developments in language, drawing, performance ability and social responses were associated with the attainment of the last sensori-motor stage, thus suggesting that the achievement of the end of the sensori-motor period is an important step in the development of severely subnormal children.

I.—INTRODUCTION.

A PREVIOUS study (Woodward, 1959) investigated a method of classifying the severely subnormal in terms of the six stages of sensori-motor development which have been distinguished by Piaget (1953 and 1955). The results indicated that severely subnormal children pass through these steps in the same sequence as Piaget suggests for normal children up to the age of two years. The aim of the present enquiry was (i) to determine the locomotor, language and social development of children classified at these sensori-motor stages, and (ii) to examine patterns of development in young severely subnormal children.

The main theoretical interest attaches to the comparison of language development and the Piaget stages. At the last sensori-motor stage, children solve problems in such a way as to suggest that they are able to anticipate the future position of an object before they move it, and they persistently search for hidden objects in a manner which suggests that they are aware of the permanence of objects. According to Piaget (1950), this behaviour indicates that the child has some means of representing objects to himself in their absence, and this implies the development of ' a symbolic function ' which makes possible the systematic acquisition of language. This suggests the hypothesis that severely subnormal children will make a similar development in language at the end of the sensori-motor period.

II.—SUBJECTS.

The subjects were resident at the Fountain Hospital, London. The total hospital population born during 1951 or later was selected in the first instance as the group to be investigated. Children were then excluded from this group if they had any condition, revealed in the medical examination, such as cerebral palsy, gross hydrocephalus, or a visual or auditory defect, that is likely to retard any aspect of development. In addition, two children with I.Qs. of over 50 were omitted from the study. This left a group of fifty-four children. During the following eighteen months twenty-nine new admissions were added to the group ; these were consecutive cases without motor or sensory disabilities. This yielded a total group of eighty-three children (fifty-one boys and thirty-two girls), aged 11 months to 8 years 7 months (mean age=4·85 years). Three children had Terman-Merrill I.Qs. of 42, 41 and 31 ; the rest failed one or more of the six two-year items on this scale (forty-nine failed all six), thus falling below the lower limit of the scale.

---

* Now at University College, Swansea.

† Now at Child Guidance Clinic, Bexleyheath, Kent.

III.—METHOD.

The main classification was made in terms of Piaget's six sensori-motor stages. Since these stages have been described in detail in the earlier study already referred to, only the main criteria are given below.

Stage I (birth to 4/5 weeks).
Use and extension of reflex activities.

Stage II (1 month to 4/5 months).
Integration of reflex activities with higher activities, e.g., co-ordination of vision and hearing and vision and grasping.

Stage III (4/5 months to 8/9 months).
Repetition in the same form of chance-made movements that produced an effect on objects that the child saw or heard (e.g., shaking a rattle). Following the line of fall when an object is dropped and recovering half-hidden objects.

Stage IV (8/9 months to 11/12 months).
At the beginning of the stage, removing an obstacle in order to grasp a toy which is visible: at the end, removing an obstacle when the object is hidden but searching only in one place. Visual examination of objects by turning them round and round and holding them at various distances.

Stage V (11/12 months to approximately 18 months).
Obtaining an object that is out of reach by pulling a string attached to it, or the support on which it stands; searching briefly in different places when an object is hidden. Exploratory manipulations which vary the effect on objects, e.g., banging a toy successively on different-sounding surfaces.

Stage VI (approximately 18 months onwards).
Inferring the result of an action before the object is moved, e.g., using a stick in order to obtain a toy that is out of reach; searching for objects for a longer period (i.e., awareness of the permanence of objects is acheived).

More advanced manipulation was investigated by the following performance tasks from the Revised Stanford-Binet (Terman and Merrill, 1937) and Merrill-Palmer (Stutsman, 1931) scales of intelligence:

Four-brick tower.
Three-figure form board.
Three-figure form board rotated.
Wall in peg boards A and B.
Block bridge.
Bead threading.
Seguin form board.
Little pink tower.
Six-cube pyramid.

If a child did not succeed with any of the first four tasks, the rest were not given.

It was subsequently found necessary to derive a performance score for each child. One point was awarded for each all-or-none task that was passed, according to Merrill-Palmer and Stanford-Binet criteria of success; timed tasks (e.g., peg boards and Seguin form board) were given weighted scores on the basis of speed. Refusals were allowed for as in the Merrill-Palmer method of scoring (Stutsman, 1931).

The subjects were in the phase of development which, in normal children, would be assessed by means of developmental scales. These were used in the present enquiry as a basis for the assessment of speech and locomotor development, but experience of administering such scales to severely subnormal children suggested certain modifications.

The locomotor section of Gesell's schedule (Gesell, 1940 and 1941) was used, with the exception of items for which special equipment, such as walking boards, was needed. Gesell's instructions for administration were followed. In cases of doubt additional information was obtained from nursing staff. At the more advanced levels, 2½ years onwards, some items (e.g., standing on one foot) require the child to imitate the examiner. The co-operation of withdrawn children could not be obtained for these items, and teaching and nursing staff could not supply information on them. Thus, these imitative items were omitted from the scoring for all children.

The language sections of developmental scales include, in addition to vocalisation and comprehension items, others relating to response to sound, manipulation of a bell and social responses. When emotionally disturbed children obtain low scores on language scales because they fail to respond to sounds, toys and people, the result does not accurately describe their language development. In this study only items of vocalisation and comprehension were included. Severely subnormal children frequently show scatter on a language scale, verbal comprehension being relatively more advanced than vocalisations. Thus, these two aspects of language were treated separately.

For speech development the vocalisation items were abstracted from Gesell's scale, together with the age norms assigned to them by him. Scores in terms of ' speech ages ' were based on these norms. Records were made of the type of vocalisations uttered by the child during the examination, and further information on habitual sounds was obtained from nursing and teaching staff. Some children made no other sounds except grunts and squeaks, and, although these occur as items in Gesell's scale, speech ages were not assigned on the basis of these alone, so that these children were omitted from the analysis of the results on speech. The amount of the child's spontaneous vocalisation during the examination was rated on a three-point scale.

The following items were used for the investigation of comprehension :

Response to a gestured instruction.
Response to simple commands in a concrete situation (e.g., " Shut the door, sit down, stand up ").
Comprehension of ' give ' in relation to a real common object (e.g., cup, spoon, ball).
Comprehension of ' show ' in relation to a real common object.
Response to the question : " Would you like a sweet ? "
Response to the statement : " There's a sweet in the box."
The simple commands item No. 6 of year II of the Terman-Merrill scale (Form L).

Data collected on social development included the child's social response to the examiner and to a familiar nurse, his response to difficulty, and any attempts at communication or at drawing the examiner's attention to his activities. Information on the achievement of feeding skills was obtained from nursing staff.

IV.—RESULTS.

(*a*) *Performance ability.*

The number of children falling in each of Piaget's sensori-motor stages is shown in Table 1, together with the average age of the children in each stage.

TABLE 1

SENSORI-MOTOR STAGE AND LOCOMOTOR AGE.

Sensori-motor Stage	N.	Mean C.A. (yrs.)	Locomotor age in years		
			Basal age Mean	Highest point	
				Mean	Range
III	10	2·48	0·43	0·60	0·38—0·77
IV	10	3·80	0·83	1·02	0·54—2·0
V	29	4·83	1·41	1·55	0·77—3·0
VI	34	5·88	2·32	2·38	0·94—4·0

Twenty-five of the thirty-four Stage VI children succeeded with two or more of the performance tasks, their scores ranging from 1 to 21, but no child below Stage VI succeeded with any of them, a difference which is significant at beyond the 0·001 level ($\chi^2=35{\cdot}4$, d.f.$=1$). Of the nine Stage VI children who failed all these items, five showed signs of severe behaviour disorder in the form of visual avoidance, or prolonged distress. The other four were the youngest at this stage, aged 3 years to 4 years 4 months.

(*b*) *Locomotor development.*

Two locomotor ages for each child derived from the Gesell scale as follows:

(1) The age level at which the child succeeded with all the items (basal age).

(2) The highest point at which a success was achieved.

The locomotor scores of children in each sensori-motor stage are shown in Table 1. From this it can be seen that the average locomotor age increases with successive sensori-motor stages, but there is a wide range at each stage. Nevertheless, the difference between the means for the highest point score of Stages V and VI are statistically significant ($t=8{\cdot}42$, $p<0{\cdot}001$, d.f.$=61$). In view of the small numbers in Stages III and IV and of the small differences between the means, these two stages were combined; the resulting average is significantly lower than the Stage V average ($t=4{\cdot}75$, $p<0{\cdot}001$, d.f.$=47$). In no instance was a child's locomotor age (highest point) level with his chronological age.

A difference between the basal age and the highest point occurred in thirty-six cases. Among the children who could not walk without support there was a consistent pattern in the type of item which accounted for the scatter. The details are given in Table 2 for the twenty-one children whose upper limit (highest point) was below one year.

TABLE 2

ANALYSIS OF ITEMS PASSED AND FAILED BY TWENTY-ONE CHILDREN SHOWING SCATTER ON THE GESELL LOCOMOTOR SCALE.

Highest point of locomotor development	N.	Type of item passed (+) and passed (—) at highest point				Range of scatter between last standing item + and last sitting item +*
		Sitting items		Standing items		
		+	—	+	—	
48 w	7	7	0	0	7	8 to 24 weeks
40 w	5	4	1	0	5	12 to 28 weeks
32–30 w	4	4	0	1	3	4 to 8 weeks
20–28 w	5	4	1	—	—	No standing items at these ages
	21	19	2	1	15	

* Age level at which last success achieved for a sitting item subtracted from age level at which last success is achieved for a standing item.

It will be noted that in nineteen cases the success at the highest point was a sitting item, and in one case only it was a standing item, and that a sitting item was failed at the highest point in only one case. Standing items, on the other hand, were failed at the highest point by fifteen of the sixteen children whose highest point was thirty-two weeks and over (there are no standing items below the age of thirty-two weeks). Thus, the scatter in the majority of cases was due to failure to show various standing responses over a range of the scale at which sitting responses were present. New admissions, who had recently come from home did not differ significantly from those who were already resident in hospital either in locomotor score or in the relative number showing a scatter over the locomotor scale.

Locomotor score (highest point) was positively correlated with performance score among the Stage VI children (rho=0·68, p=0·001). Hence, this factor was controlled in group comparisons, as was chronological age also, though a significant correlation between age and locomotor score (highest point) within any stage was found only for the group at Stage VI with performance scores of 0 (rho=0·625).

In view of the tendency of Mongols to lower muscle tone (Hilliard and Kirman, 1957), their locomotor development was compared with that of non-Mongols. Twenty-three pairs were matched for age, Piaget stage and performance score, the average ages of the two groups being 4·96 years and 5·025 years, respectively. The mean locomotor age (highest point) of the Mongols which was 1·76 years (S.D.=0·98) differed very little from that of the non-Mongols (M=1·77, S.D.=0·96). The basal age means were similarly close (1·63 years and 1·64 years).

(*c*) *Speech development.*

In fifteen cases no sounds (apart from squeaks, crying and laughing) were heard during the examination, and the nursing staff similarly reported having heard no other sounds from these children. The ' speech age ' of the remaining sixty-eight children is shown in Table 3.

TABLE 3

LOCOMOTOR AND SPEECH DEVELOPMENT.

Sensori-motor Stage	N	Locomotor age Highest point Mean (years)	Speech Age	
			Mean (years)	Range (years)
III	6	0·617	0·487	0·23—0·77
IV	8	1·058	0·403	0·15—0·69
V	21	1·660	0·784	0·54—1·25
VI	33	2·330	1·196	0·62—2·0

The first interesting feature of this table is the wide range in speech age at any Piaget stage ; the second is the increase in the upper limit of speech development with successive stages. At Stages III and IV the most complex sound made by any child was a four-syllable babble ; jargon, interwoven with single words, was the most advanced development at Stage V. Two Stage V children produced two-word combinations in an echolalic manner, but meaningful and relevant word combinations and short sentences occurred only in the speech of some of the Stage VI children.

The difference between the means of the Stage V and Stage VI children is statistically significant at beyond the 0·001 level ($t=3{\cdot}92$ for 52 d.f.). The Stage V average is also significantly different from the combined mean of Stages IV and III ($t=3{\cdot}52$, $0{\cdot}01>p>0{\cdot}001$, d.f.$=33$).

In view of the mutism that is sometimes associated with severe behaviour disorders, it was postulated that the absence of speech sounds, or very retarded speech development, would be related to behaviour disturbance in children at Stages V and VI, among whom absence of speech could scarcely be attributed merely to extremely slow development. Very retarded speech was defined as up to and including twenty-eight weeks (vowels only) for Stage V and up to and including thirty-six weeks (two-syllable babble) for Stage VI. Behaviour disorder was defined by marked visual avoidance or prolonged distress when approached, or mannerisms which warded off social contact. Eight of the eighteen children at Stages V and VI who were extremely retarded in speech development, or who made no sounds, showed signs of behaviour disorder, compared with three of the remaining forty-five children at these stages ; this difference is significant between the 0·01 and 0·001 levels ($\chi^2=10{\cdot}25$, with Yate's correction applied, d.f.$=1$).

Among the children at Stage VI, speech development was positively correlated with performance score when the four children with severe behaviour disorders were omitted (rho$=0{\cdot}38$, $p=0{\cdot}05$). Rank order correlations of age and speech development within each Piaget stage and performance score group were not significantly different from zero in any instance.

The bearing of the period of residence in hospital on speech development was examined by comparing the speech development of twenty-three new admissions and twenty-three others matched for age, Piaget stage and performance score where applicable. The average speech ages of the two groups were 0·83 years and 0·81 years, respectively. (When new admissions were excluded, there were too few children left at any stage for length of stay in hospital to be correlated with speech age.)

Twenty-two pairs of Mongols and non-Mongols were matched for age, Piaget stage and performance score. A mean difference of 0·126 years in speech age, favouring the Mongols, was found, but there were six children with severe behaviour disorders in the non-Mongol group and non among the Mongols. Excluding these six children and their Mongol pairs, the mean speech ages for the remaining sixteen pairs were 1·028 years (S.D.=0·405) for the Mongols and 0·896 (S.D.=0·5) for the non-Mongols ; this difference was not statistically significant.

A comparison of speech and locomotor development in the same children within each Piaget stage is made in Table 3, excluding the children who made no speech sounds. It will be seen that, for all groups, the average locomotor age is considerably in advance of the average speech age.

For the whole group of sixty-eight children, the average speech age (0·92 years) was significantly lower than the average locomotor age (1·82 years) for the highest point score. (t for correlated means=9·82, $p < 0{\cdot}001$, d.f.=67).

In fifty-five of the sixty-eight cases (81 per cent.), both the highest point of locomotor development and the basal age were higher than the speech age. In four cases one or both was the same as the speech age ; both locomotor ages were lower than the speech age in four cases, and in the remaining five cases only the basal age was lower.

(*d*) *Verbal comprehension.*

None of the children classified at Stages III and IV showed any understanding of speech, nor did twenty-five of those at Stage V, nine of whom responded to instructions given by gesture. The remaining four Stage V children gave evidence of verbal comprehension in simple concrete situations. All the children classified at Stage VI, on the other hand, showed evidence of some verbal comprehension. Fourteen of them responded only at the level of simple concrete situations ; of these nine failed the other items listed in Section III, but the other five were too disturbed for these to be administered. The remaining twenty children succeeded with three or more of the items listed in Section III (including concrete situations and excluding gestures). Comparing Stages V and VI for the number of children in three categories of verbal comprehension (1) success with three or more items, (2) verbal comprehension in concrete situations only, and (3) none at all, the differences are significant at beyond the 0·001 level (χ^2=50·36, d.f.=2).

The relation between verbal comprehension and Piaget's sensori-motor stages was further investigated by an examination of clinical records, which contained the Piaget stage and the results of attempts to administer the Terman-Merrill revision of the Binet scale. Children who succeeded with one or more of the year II verbal items were invariably at sensori-motor Stage VI.

(*e*) *Drawing.*

The drawing ability of children at the various Piaget stages is shown in Table 4.

TABLE 4

SENSORI-MOTOR STAGE AND DRAWING.

Drawing	Sensori-motor stage			
	III	IV	V	VI
Nothing	10	10	20	0
Faint marks or dots	—	—	7	3
Straight scribble	—	—	2	13
Circular scribble	—	—	—	3
Copy of lines	—	—	—	6
Copy of circle, partial*	—	—	—	6
Copy of circle, correct	—	—	—	2
Copy of Cross	—	—	—	0
Totals	10	10	29	33

* By circular scribble.

It will be seen that the use of a pencil begins to occur at Stage V, but extends to vigorous scribbling in only two cases. All but three of the Stage VI children, on the other hand, scribbled freely or copied models of straight lines and circles. (One refused to handle the pencil.) The differences between Stages V and VI, in three drawing categories ((1) none, (2) marks or dots, and (3) scribbles or copies) are significant at beyond the 0·001 level ($x^2=28{\cdot}1$, d.f.$=2$).

(*f*) *Feeding skills.*

Feeding skills achieved by children at various Piaget stages are shown in Table 5.

TABLE 5

SENSORI-MOTOR STAGE AND FEEDING SKILLS.

	Sensori-motor Stage			
	III	IV	V	VI
Cup · Does not drink	10	7	7	—
Drinks if cup held	—	3	8	—
Helps to hold cup	—	—	5	2
Holds cup without help	—	—	9	32
Does not feed self with fingers	10	7	13	—
Feeds self with fingers	—	3	16	34
Spoon : Does not use	10	10	23	2
Uses and spills	—	—	5	6
Uses without spilling	—	—	1	26
Total in each stage	10	10	29	34

(*g*) *Social responses.*

All the children in the group showed awareness of the investigator and a nurse or teacher, most by looking at the investigator when she approached; the social awareness of some was indicated in a negative way by visual avoidance. Others were distressed in the presence of the investigator, or began a mannerism or slapped their faces when approached.

The response of smiling was found at all stages, though not all the children in a given stage showed it; the incidence was naturally affected by the behaviour described in the previous paragraph. Few children made a vocal response when talked to by the investigator, but spontaneous vocalisations occurred in more cases. A rating of 1 (few or no spontaneous vocalisations) compared with a rating of 2 or 3, was given to significantly more of the children who showed signs of behaviour disorder than to the rest ($\chi^2=20{\cdot}02$, p. $<0{\cdot}001$, d.f.$=2$.).

Differences in the amount of vocalisation (assessed by the ratings) between new admissions and children already resident in hospital were not statistically significant.

Communication, in terms of imparting information or making a demand, was observed mainly among the Stage VI children, twenty-two of the thirty-four pointing for this purpose, with or without a vocal accompaniment. Only two Stage V children communicated, and none below this level did so.

Other social responses were observed only among the Stage VI children. Seventeen (50 per cent.) of them drew the investigator's attention to their activities, or kept looking to make sure the investigator was watching their performance; and twelve children sought the investigator's help when in difficulty.

V.—DISCUSSION AND CONCLUSIONS.

There are two main questions arising from these findings that call for discussion. One concerns the pattern of development in severely subnormal children, and the other is the relation of other aspects of development to Piaget's sensori-motor stages.

The group is too deviant for locomotor and speech scores to be equated in terms of standard scores (in any case Gesell does not provide the necessary data for this). The expression of the results in terms of ' locomotor ages ' and ' speech ages ' is probably inexact, and consequently comparisons of rates of development in different aspects are approximate. Nevertheless, the difference between the locomotor score and the speech score of the whole group is so large that it would appear that the severely subnormal, although grossly retarded in all areas of development, tend to develop much more rapidly in the locomotor sphere than they do in the acquisition of speech. If the approximate ages suggested by Piaget for the sensori-motor stages are accepted, then at all stages (with the possible exception of III) locomotor development appears to be in advance of the development of sensori-motor intelligence, which, in turn, is in advance of the development of speech.

The twin study of Gesell and Thompson (1929) indicates that locomotor development is mainly influenced by maturation, since intensive training made little difference to subsequent locomotor scores. Much of Piaget's account of the development of sensori-motor intelligence is a description of a learning process, and learning obviously enters into the acquisition of speech. Thus, it can be postulated from these results that the severely subnormal develop relatively more rapidly in the aspect which is mainly maturational, and that they develop relatively more slowly in aspects which depend more on learning.

Another finding on locomotor development which calls for further investigation is that of a large scatter over the locomotor scale by the majority of the children who could not walk. This suggests that steps in the sequence of behaviour that lead to the attainment of the upright posture do not, in the severely subnormal, correspond with steps in the development of the sitting posture as in normal children. One possible psychological factor accounting for this is a difference in the amount of individual attention and encouragement to stand received by children in hospital and at home. New admissions did not, however, differ significantly from the children who had been resident in the hospital for a longer period, either in locomotor score, or in scatter over the locomotor scale. Although children with cerebral palsy were excluded from the study, it is, of course, possible that there were motor lesions which were not observable clinically. Comparison of these data with post-mortem findings would be valuable.

The limited verbal stimulation of children in hospital might also be put forward to account for the finding of greater retardation in speech development, relative to locomotor development and the sensori-motor stages. The results of a number of studies (reviewed by McCarthy, 1954) have indicated that normal children living in institutions develop language more slowly than children living at home. A similar finding in respect of severely subnormal children has been reported by Lyle (1960). However, in the present study new admissions did not differ in average speech score from children already resident in hospital. This finding is not necessarily at variance with Lyle's, since he investigated older children whose language was more advanced. They could identify and name objects, whereas the present study was mainly of pre-verbal utterances. It is thus possible that different types of environment become more important once words are beginning to be systematically acquired. The studies of normal children, however, give conflicting results as to the age at which an unstimulating environment begins to affect language development, some reporting an effect on phonemes in the first few months and others finding a difference only after the age of 2 years (McCarthy, 1954). The hypothesis that an unstimulating environment influences language development only after the beginning of the systematic use of speech might also account for the different results in respect of Mongols in this study and in Lyle's. Lyle (1960) found that Mongols, whether living in hospital or at home, were more backward in language than comparable non-Mongols. The Mongols of the present group did not differ significantly from non-Mongols.

Further study might usefully be made of the findings of the present study that children with severe behaviour disorders tended to be very retarded in speech development or to make no sounds. This result also indicates that behaviour disorder is a variable that should be controlled in comparative studies of language development.

The second main question to be discussed is whether children classified into Piaget's sensori-motor stages are clearly differentiated in other aspects of behaviour. For most aspects of development there is little difference between Stages III and IV. This is probably to be explained by the fact that all the children at Stage IV were only at the beginning of this stage. This appears to be a transitional stage, and very few of the severely subnormal children who have been observed have been at the end of the stage. Children are classified at the beginning of Stage IV if they brush aside a transparent screen in order to obtain a toy that it covers, though in the extent of their awareness of the permanence of objects they are the same as those at Stage III (recovering a half-hidden object, but not a completely hidden one).

However, significant differences in both locomotor and speech development were found between successive stages (III and IV together compared with V, and V compared with VI). Children at Stage VI were significantly differentiated from those at Stage V in the development of verbal comprehension. It is also of particular interest that it is only Stage VI children who can succeed with, or even comprehend, the first verbal items of the Revised Stanford-Binet scale. For these it is necessary to identify objects by their names and to carry out simple instructions about objects. This finding supports Piaget's contention that the achievement of Stage VI, which involves a concept of the permanence of objects, is necessary for the acquisition of meaningful speech. Some children in the present group who were at Stage VI did not use meaningful speech, or identify objects by name; whether these only lacked sufficient stimulation, or whether another factor is involved, as a matter for further investigation.

It was also only Stage VI children who succeeded with any performance items drawn from pre-school scales. Thus, the attainment of Piaget's sensori-motor Stage VI, at any rate in the severely subnormal, appears to be necessary for success with the earliest items, both verbal and non-verbal, of pre-school scales which are applicable to normal children from the age of 1½ to 2 years.

Results for other aspects of behaviour also underline the importance of the attainment of the end of the sensori-motor period in the development of severely subnormal children; only Stage VI children were able to copy lines and a circle, and with two exceptions vigorous scribbling occurred only at this stage. Communication began to appear at Stage V but was common only at Stage VI, and it was only Stage VI children who sought help in difficulty, and who indicated that they wished their activities and achievements to be noticed by an adult.

To turn to the practical implications of the findings, the first application is in clinical practice. All the Stage VI children in this study showed evidence of some verbal comprehension. Hence, if a child at Stage VI is found to have no verbal comprehension, then further investigation is indicated, particularly of the possibility of deafness.

The results of this study also suggest when it is useful for nursing staff to spend time encouraging a child to use a cup and spoon. The findings indicate that children who have reached Stage VI can certainly be taught to feed themselves, and they also suggest that it is worthwhile to persist with Stage V children. However, the attempt is probably fruitless until the child reaches these two stages.

In the educational field the results indicate that children who have attained the last sensori-motor stage would benefit from attending a training centre, since new developments in language, drawing and performance ability occur during this stage, which appears from this study to be a major event in the development of severely subnormal children.

VI.—REFERENCES.

GESELL, A. (1940). *The First Five Years of Life.* New York: Harper.

GESELL, A., and THOMPSON, H. (1929). Learning and growth in identical infant twins. *Genet. Psychol. Monag.*, 6, 1-124.

GESELL, A, and AMATRUDA, C. S (1941). *Developmental Diagnosis.* New York and London: Hoeber.

HILLIARD, L. T., and KIRMAN, B. H. (1957). *Mental Deficiency*. London : Churchill.

LYLE, J. (1960). The effect of an institution environment upon the mental development of imbecile children : II—Speech and language. *J. Ment. Defic. Res.*, 4, 1-13.

McCARTHY, D. (1954). Language development in children. Carmichael, L. (Ed.). *Manual of Child Psychology*. New York : John Wiley.

PIAGET, J. (1950). *The Psychology of Intelligence*. Trans. by Piercy, M., and Berlyle, D.E. London : Routledge and Kegan Paul.

PIAGET, J. (1953). *The Origin of Intelligence in the Child*. Trans. by Cook, M. London : Routledge and Kegan Paul.

PIAGET, J. (1955). *The Child's Construction of Reality*. Trans. by Cook, M. London : Routledge and Kegan Paul.

STUTSMAN, R. (1931). *Mental Measurement of Pre-School Children*. Yonkers, New York : World Book Co.

TERMAN, L. M., and MERRILL, M. A. (1937). *Measuring Intelligence*. London . Harrap.

WOODWARD, M. (1959). The behaviour of idiots interpreted by Piaget's theory of sensorimotor development *Brit. J. Educ. Psychol.*, 29, 60-71.

(*Manuscript received*, 17*th May*, 1962)

2H3679

ARE RURAL CHILDREN HANDICAPPED BY THE USE OF SPEEDED TESTS IN SELECTION PROCEDURES ?

BY
C. ANNE MORETON
(*Charlotte Mason Training College*)
AND
H. J. BUTCHER
(*Department of Education, University of Manchester*)

SUMMARY. It is widely believed in country areas that the country child is at a disadvantage in the working of speeded tests. This hypothesis was tested by the comparison of two samples, each of about 300 children aged 10–11, one from rural Westmorland, one from central Manchester. Power and speed scores were obtained from N.F.E.R. tests of Arithmetic, English and Verbal Reasoning Almost all net differences were statistically significant in the expected direction, tending to confirm the hypothesis, but were not sufficiently large to suggest a need for revision of selection procedures.

I.—INTRODUCTION.

IN the course both of teaching in rural schools and also of supervising school practice in a rural area, one of the authors has frequently met with complaints from teachers and parents that the country child is at a disadvantage in the working of speeded tests used in selection procedures. This disadvantage is often attributed to the slower tempo of life in the country as compared with that in urban areas.

The present research is confined to testing the hypothesis that such a disadvantage exists, without any detailed exploration of possible reasons or correlates. But it is believed to be the first systematic enquiry of its kind and may be timely in view of recent renewed interest in the concepts of speed and power in mental tasks (e.g., Furneaux, 1960).

II.—PREVIOUS RESEARCH.

The literature on mental 'speed' and 'power' is extensive, and often conflicting. About 150 references are given by Morrison (1960) and it is, therefore, outside the scope of this paper to offer anything like a comprehensive survey. One of the difficulties in this field of research is that in practice measures of both speed and power are usually adulterated or relative, since a pure test of speed is of trivial application, and a pure test of power impracticable. Furneaux (1960) is probably correct in holding that such concepts as 'speed,' 'power,' 'accuracy' and 'difficulty' require a new and more operational definition if conflicting results are to be avoided. However, certain trends—one might almost say fashions—in research do appear to emerge.

In the earliest researches there was disagreement between those who maintained that speed of reaction was positively correlated with higher mental processes (Gilbert, 1894, Baguley, 1901) and those who found no such relationship (Binet, 1900 ; Wissler, 1901 ; Whipple, 1904). As a result of further research (Burt, 1909 ; Brown, 1910 ; Kruger and Spearman, 1914 ; May, 1921 ; Bernstein, 1924; Spearman, 1927), the view that such a relationship does exist became generally accepted, although the relative importance and predictive power of these two components of mental ability is still not entirely settled.

Since about 1930, attempts have been made, foreshadowed by Bernstein (1924), to interpret speed of mental work as a personality factor (e.g., Kennedy, 1930; Himmelweit, 1946). In contrast, Eysenck (1953), discussing the work of Furneaux, has maintained that speed is the fundamental component of ability, and that measures of power, which depend on persistence, involve personality factors.

Broadly speaking, however, the orthodox view appears still to be, firstly that speed and power are highly correlated, secondly that power is an important component, if not the more important component of general ability, thirdly, that tests, as used for example in the 11+ selection procedure, achieve an adequate compromise. Thus, Peel (1956), commenting on the kind of criticims that gave rise to the present investigation, concludes both that in general at 11+ " the abler child is the speedier child," and that in any case the time-limit imposed on most group tests in practice is not so stringent as to make speed an over-riding factor.

Among psychologists as well as among the general public there are some who might still query this view (Porebski, 1954; Myers, 1952, 1960; Mollenkopf, 1960). Myers, for example, writes: " In timed selection and achievement tests, we may be defining the purposes of education as including, perhaps emphasising, the teaching of the value of snap judgment and educated guessing, over the value of deliberate careful thought."

In view of the continuing interest in these questions, it is mildly surprising to find how little research has been done into the social correlates of power and speed. Klineberg (1928), Lambeth and Lanier (1933) and Knapp (1960) investigated racial differences, but urban-rural differences appear to be relatively unexplored. The most relevant study known to the writers is that of Smith (1948), whose conclusions about Hebridean children include the following:

> " It is clear in these circumstances that the time factor has a profound effect on test results over and above the normal effect. A rigidly timed test does not seem to discriminate very well in the range of the intelligence scale, because mental processes which are more slowly initiated and pursued than is usual are constantly being interrupted and suspended by the imposed speed of the examiner . . . There is little evidence of the modern tendency to race with time. It is rightly considered infinite and unlikely to interfere with the expected sequence of things; so in the normal course of events there is no need for undue haste. Probably if the children had grasped more fully the significance of the stop-watch in the experiment, the scores would have been more flattering to them."

Smith's research was primarily concerned with bilingual children, and there can be few rural areas comparable in isolation with the Outer Hebrides, yet it seems reasonable to postulate a gradation of natural tempo between such children and urban children. The aim of the present research is to determine whether English rural children in a relatively isolated area are handicapped to some degree in a similar way.

III.—CHOICE OF SAMPLES.

(*a*) *The Manchester sample.*

The Manchester sample of 297 was drawn from six schools in a lower working-class area near the centre of the city. As will be seen from Table 1, the largest number from any one school was 74 and the smallest number 36. In the total sample there were 153 boys and 144 girls. The average age was 10·11.

TABLE 1

School	No. on Roll	Girls	Boys	Total
1	254	17	19	36
2	247	28	32	60
3	228	21	19	40
4	302	40	34	74
5	290	27	34	61
6	271	11	15	26
			N=297	

School No. 1 was a junior mixed school on the edge of a rapidly deteriorating area and, according to the headmistress, has over the past five years lost a number of good working-class families who have moved out to one of the newer estates. Two coloured boys were concerned in the tests, a half-caste and a Nigerian; many coloured families are now invading this part of Manchester. There were several 'problem' families, but on the other hand, still many families who had lived in the district for one or two generations.

School No. 2 was a junior and infant school Again, there was some evidence of deterioration, both to property and living standards in this area, although on the whole, the school had maintained its standards. There were several coloured children in the school who seemed to be at least of average ability and who also seemed to fit in very well with the community.

School No. 3 was a junior school. It was in a similar type of area to that of School No. 2, although here there was more evidence of deterioration of property and living standards. The coloured children did not appear to be of the same calibre as those in the previous school.

School No. 4 was a junior school. This school was in rather a better residential working-class area but the population of the school was a little mixed and there was some slight evidence of a deterioration in the background and living standards of the children over the past few years, although actually the proportion of those selected for grammar schools was above average.

School No. 5 was what is generally known as a slum school. There is a large floating population and many families live in rooms and sub-standard houses. There are about thirty coloured children in the school. It is interesting to note that at this school a coloured boy who took no extra time for any of the tests had both the highest speed raw score and quotient and also the highest time-limit raw score and quotient of the Manchester sample in Arithmetic.

School No. 6 was the nearest to the centre of the city. Some of the children were from very good working-class homes with interested parents, others were from poor homes both economically and socially. There was a small proportion of coloured children in the school and also a variety of nationalities (e.g., Spaniards, Nigerians and Indians) with attendant language difficulties. Considering the prevailing living conditions, the number going forward to either grammar or technical schools or grammar-stream in secondary modern schools was relatively high.

(*b*) *The Westmorland Sample.*

The Westmorland sample of 274 was drawn from all the rural schools in the county except two which had to be withdrawn owing to administrative difficulties. The largest number from any one school was twelve and the smallest number one. There were 142 boys and 132 girls in this sample, with an average age of 10·8. The total number of schools involved was fifty-eight, the largest number on roll being ninety-six, and the smallest six. There were twenty-five one-teacher schools, twenty-seven two-teacher schools, five three-teacher schools and one four-teacher school. It is probably true to say that the majority of the children in these schools were the children of agricultural workers, although there is a slight but steady trend towards an increasing proportion of children whose fathers work in the towns and travel out each day. The majority of such parents are office workers and/or members of the professions.

Considerable care was taken to confine the sample to genuinely rural areas, by the omission of even small towns and of areas that are fast becoming urbanised. The number of old Westmorland family names gives support to the assumption of a strong ' native ' element. There are for example, thirteen of these known to the writers, all of which occur on an average four times. It is not claimed, however, that the two samples are equated for socio-economic status, and it is probable that the Westmorland sample contained a higher proportion of children from middle-class homes.

IV.—METHOD.

(*a*) *The tests used and their administration.*

The tests used were the NFER Arithmetic Test 8, English Test 8, and Verbal Reasoning Test 8a. Considerable practical difficulties were envisaged in conducting a programme of testing over a wide rural area with poor communications and a high proportion of one-teacher schools, in which such a programme can easily disrupt the whole school.

The solution adopted was that the Westmorland results should be based on the Authority's preliminary testing, with some additional instructions to supervisors, and that the same tests should be administered to the Manchester children under as closely similar conditions as possible (but without, of course, forming part of the Manchester Authority's official programme of testing).

Such a solution is not ideal, and might even introduce differences in absolute level of performance, but it was considered unlikely seriously to disturb the *relative* differences between speed and power which form the main topic of investigation.

Instructions to Supervisors.

In addition to the usual instructions for administration issued with the tests as to strict time limit, supervision and marking, the following instructions were sent to each school:

Additional Instructions to Supervisors
Arithmetic Test 8 (50 *Minutes*), *English Test* 8 (50 *Minutes*)
Verbal Reasoning Test 8a (45 *Minutes*)

(1) After each test has been in progress for 10 minutes, instruct the testees to place an 'X' in the margin.

(2) Record time taken by each testee to complete test within the official time-limit. This to be recorded on list containing names and ages (in years and months).

(3) At the earliest convenient time (possibly the following day) allow testees to have back their tests if they are not complete, and to complete them in red pencil. (N.B.—All items omitted on the previous occasion must be completed.)

(*b*) *Types of Score desired.*

The ten-minute score is referred to as a *Time-Limit Score* following Davidson and Carroll's (1945) definition as "the number of items correctly answered within a prescribed time-limit." The score achieved on the normal working of the test, i.e., on the first day and within the 'official' time is referred to as the *Speed Score*. The distinction between 'time-limit score' and 'speed score' is thus one of degree rather than of kind. Finally, the total score including any red pencil entries, is called the *Power Score*. Thus, the experiment was designed to give nine scores in all for each child, consisting of three types of score on each of three tests.

(*c*) *Method of standardisation.*

In order to test our hypothesis, it was necessary to make power scores, time-limit scores and speed scores all comparable. Conversion tables incorporating an age allowance and giving quotients were available in the case of speed scores, but not, of course, in the case of the other two types of score, which were peculiar to the experiment.

The power and time-limit scores were, therefore, scaled to the same mean and standard deviation as the speed scores for the combined sample. It was then possible to convert these into quotients comparable with the speed quotients. A section of one of the conversion tables used is reproduced in Table 2.

TABLE 2

SECTION OF CONVERSION TABLE (ARITHMETIC TEST)
(Ages 10 years—10·11 years)

Stand. Score	Raw Score	Age in years and completed months at time of test												Raw Score
		10·0	10·1	10·2	10·3	10·4	10·5	10·6	10·7	10·8	10·9	10·10	10·11	
−1·723	1	79	78	77	77	76	76	75	75	74	74	73	73	1
−1·678	2	82	81	81	80	80	79	79	78	77	77	76	76	2
−1·571	3	84	84	83	83	82	81	81	80	80	79	79	78	3
−1·588	4	86	86	85	84	84	83	83	82	82	81	81	80	4
—1·544	5	88	87	86	86	85	85	84	84	83	83	82	81	5

V.—RESULTS.

(*a*) *A direct comparison of the two samples on speed, power and time-limit scores.* Table 3 summarises the comparative performance of the two groups on the three tests and under the three conditions of work. The power and time-limit scores have been made comparable with the speed scores by the procedure already described, which also incorporates an age allowance :

TABLE 3

MEANS AND STANDARD DEVIATIONS—CONVERTED SCORES.

	MANCHESTER		WESTMORLAND		Mean Diff.	Crit. Ratio	Level of Signif.
	Means	S.D.	Means	S.D.			
ARITHMETIC							
Speed	105·32	13·49	104·83	12·21	+0·49	·3571	N/S
Power	104·42	12·9	104·55	12·99	−0·13	−·1199	N/S
Time-limit	106·29	13·71	103·17	13·33	+3·12	2·756	1%
ENGLISH							
	MANCHESTER		WESTMORLAND				
Speed	103·94	13·47	101·82	11·14	+2·12	2·055	5%
Power	103·36	13·11	102·18	12·04	+1·18	1·222	N/S
Time-limit	104·16	13·19	101·87	12·02	+2·29	2·171	5%
VERBAL REASONING							
	MANCHESTER		WESTMORLAND				
Speed	104·44	12·40	102·16	11·78	+2·28	2·255	5%
Power	104·07	12·44	102·58	12·12	+1·49	1·450	N/S
Time-limit	104·34	12·25	100·65	13·17	+3·69	3·426	1%

It can be seen from Table 3 that any significant differences are in favour of Manchester group. Furthermore, all these differences are in respect of the speeded conditions of work (i.e., ' speed ' and ' time-limit '). Thirdly, in each test, the more highly speeded the conditions the greater or more significant the difference in favour of the urban children. This is particularly noticeable in the case of arithmetic, but applies in some degree to the other two tests as well.

(*b*) *A comparison of the two samples on difference scores (e.g., power minus speed).*

The results reported in the last section already suggested that the urban children were superior to the rural children to a markedly greater degree in the speeded than in the unspeeded conditions, although absolute differences were small. But a crucial test of the main hypothesis depends on a study of second-order differences, i.e., a comparison of the urban and rural children on power-score minus speed-score.

Accordingly, the difference scores were tabulated and analysed, with results as shown in Table 4 below.

TABLE 4

COMPARISON OF WESTMORLAND AND MANCHESTER CHILDREN ON POWER SCORE MINUS SPEED SCORE.

	Manchester	Westmorland	Difference	Critical Ratio	Significance Level
Arithmetic	−0·90	−0·28	−0·62	2·17	·05
English	−0·58	0·36	−0·94	3·30	01
Verbal Reasoning	−0·37	0·42	−0·79	2·89	·01

It can be seen from Table 4 above that all three crucial differences, although small, are in the expected direction and are statistically significant. The size of the difference is largest for English and smallest for arithmetic, although the range is not very great, and should in any case not be interpreted too strictly, since the standardising procedure in the experiment ensured comparability of speed and power scores for any one test, but not necessarily between tests.

A similar comparison in respect of power scores minus time-limit scores is shown below in Table 5 .

TABLE 5

COMPARISON OF WESTMORLAND AND MANCHESTER CHILDREN ON POWER SCORE MINUS TIME-LIMIT SCORE

	Manchester	Westmorland	Difference	Critical Ratio	Significant Level
Arithmetic .	−1·87	+1·38	−3·25	5·30	·01
English .	−0·80	+0 31	−1·11	1·82	n s
Verbal Reasoning	−0·27	+1·93	−2·20	4·03	·01

The second-order differences shown in Table 5 are again all in the hypothesised direction, two being highly significant, and the third close to statistical significance. Since the ' time-limit ' scores are formally similar to ' speed ' scores, but based on a shorter period of testing time, it might have been supposed that less conclusive results would be obtained than in the comparison between power scores and speed scores. The indication is, however, that even within the short space of ten minutes the measurement of speeded work operates to some extent in favour of urban children.

VI.—CONCLUSIONS.

1.—Children from a working-class area of central Manchester were compared with children from a rural area of Westmorland on moderately speeded tests of arithmetic, English and verbal reasoning as used in an 11+ selection procedure. Manchester scores were slightly higher on the average in all three papers, two of the three differences being statistically significant.

2.—Scores were obtained after 10 minutes' work on the forty-five minute or fifty-minute tests. Rather larger differences were found in favour of the Manchester children.

3.—When the children were allowed unlimited time to complete the papers, the mean differences between the groups were reduced, and none was significant.

4.—Power-minus-speed scores were calculated for each child. Westmorland children were slightly but significantly higher on all three papers.

5.—The hypothesis that speeded tests to some degree handicap rural children was considered to be confirmed. Relative differences, however, being of the order of one unit of I.Q., were so small as not to suggest any revision of the selection procedure.

VII.—REFERENCES.

Baguley, W. C. (1901). On the correlation of mental and motor ability in school children. *Am. J. Psychol.*, 12, 193-205.

Bernstein, E. (1924). Quickness and intelligence. *Brit. J. Psychol. Monogr. Supp.* 3, No. 7.

Binet, A., and Henry, V. (1896). La psychologie individuelle. *L'Année Psychol.*, 2, 119.

Brown, W. (1910). Some experimental results in the correlation of mental activities. *Brit. J. Psychol.*, 3, 296-322.

Burt, C. (1909). Experimental tests of general intelligence. *Brit. J. Psychol.*, 3, 94-177.

Carroll, J. B., and Davidson, W. M. (1945). Speed and level components in time-limit scores : a factor analysis. *Educ. Psychol., Measmt*, 5, 411-427.

Eysenck, H. J. (1947). Dimensions of personality, Chapter 4, Section 5. London : Kegan Paul.

Eysenck, H J. (1953). *Uses and Abuses of Psychology*. 30-37, Pelican.

Furneaux, W. D (1960) Intellectual abilities and problem-solving behaviour, Chapter 5 in *Handbook of Abnormal Psychology*, H J. Eysenck, Ed.

Gilbert, T. A. (1894). Researches on the mental and physical development of school children. *Yale Psychology Laboratory*, 240-300.

Himmelweit, H. (1946). Speed and accuracy of work as related to temperament. *Brit. J. Psychol.*, 36, 132-144.

Kennedy, M. (1930). Speed as a personality trait. *J. Soc. Psychol.*, 1, 286-299.

Knapp, R. R. (1960). The effects of time limits on the intelligence test performance of Mexican and American subjects. *J. Educ. Psychol.*, 51, 14-20.

Klineberg, O. (1928). An experimental study of speed and other factors in " racial " differences. *Arch. Psychol.*, No. 241.

Lambeth, N., and Lanier, L H. (1933). Race differences in speed of reaction *Pedagog. Sem.*, 42, 255-297.

May, M A. (1921). Psychological examination in the U.S. Army. *National Academy of Sciences Memoirs*, 15, 415-420.

Mollenkopf, W. G. (1960). Time limits and the behaviour of test takers. *Educ. Psychol. Measmt.*, 20, 223-230.

Morrison, E. J. (1960). On test variance and the dimensions of the measurement situation. *Educ. Psychol, Measmt*, 20, 231-250.

Myers, C. T., (1952). The factorial composition and validity of differently speeded tests. *Psychometrika*, 17, 347-352.

Myers, C. T. (1960). Introduction to symposium on the effects of time limits on test scores. *Educ. Psychol Measmt.*, 20, 221-222.

Peel, E. A. (1956). *The Psychological Basis of Education*, 132-4. Edinburgh : Oliver and Boyd.

PORBBSKI, O. (1954). A pyschological and statistical study of speed and power as variables of human ability. *Occup. Psychol.*, 28, 218-229.
SMITH, C. A. (1948). *Mental Testing of Hebridean Children in Gaelic and English.* London : University of London Press.
SPEARMAN, C. (1927). *The Abilities of Man.* London : Macmillan.
SPEARMAN, C., and JONES, W. (1950). *Human Ability.* London : Macmillan.
WISSLER, C. (1901). The correlation of mental and phsycial tests. *Psychol. Monogr.*, 1662.
WHIPPLE, C. M. (1904). Reaction time as a test of mental ability. *Am. J. Psychol.*, 15, 489-498.

(*Manuscript received*, 23*rd May*, 1962.)

SOME CHARACTERISTICS OF CHILDREN WITH TWO TYPES OF SPEECH DISORDER

BY
T. G. CROOKES
(St. John's Hospital, Aylesbury)
AND
MARGARET C. L. GREENE
(Stoke Mandeville Hospital, Aylesbury)

SUMMARY. An investigation was made of twenty successive cases of developmental speech defects. They were divided according to the type of speech disorder (Greene 1960) into a ' Motor ' group (eight children) and a ' Language ' group (twelve children). In the Motor group, but not the Language, there was evidence of delay in general motor maturation. The two groups were compared on measures of mechanical reading, consistency of laterality and W.I.S.C. subtests. Both groups were very poor at reading, but whereas all the Motor group had inconsistent laterality, only a third (i.e., a normal proportion) of the Language were inconsistent; on the other hand, the Motor group had a much better knowledge of right and left. The Language group, but not the Motor, had poor scores on the Arithmetic and Digit Span subtests. Also the Language group was significantly worse than the Motor on Block Designs.

It was suggested that these might represent two distinct groups of developmental language defects, and that recognition of the two might help to explain some of the contradictions in the literature.

I.—CLINICAL BACKGROUND.

THIS is a preliminary attempt to measure by psychological tests certain clinically observed features of children suffering from dyslalia and referred to the speech therapist between the ages of 3–6 years. It was possible for the children to be sorted into two groups on the basis of speech performance alone as already described by Greene (1960) and classified as suffering from a ' motor ' or ' language ' disorder. The speech of the youngest children at first exhibited characteristics of both motor and language disorder. This was to be expected since the normal child has the same difficulties in handling articulation and language when learning to speak, though to a lesser degree. By the age of 4 : 0–4 : 6, the pattern of a predominantly motor disorder on the one hand and a language disorder on the other had emerged. Diagnosis in each case was supported by a careful case history of general and speech development and the tests for motor and language maturation detailed below.

Motor Disorder Group.

The children classified as suffering from motor disorder had great difficulty with articulation and the accurate imitation of consonants and vowels, especially complex consonant blends such as sp, sk, str, etc. The omission and substitution of consonants followed a pattern which reduced the conventional phonetic sequences of the language to simpler forms, for example ' biscuit ' pronounced ' bitit,' ' swim ' as ' fim.' Diphthongs were frequently reduced to vowels as when ' boy ' was pronounced ' bore.' These mispronunciations were consistent as when ' sk ' was invariably substituted by ' t ' and ' sw ' by ' f.' Once the child's individual code was known, it was possible to understand what he said. Vocabulary and language development were normal for age and the Watts picture tests presented no difficulty. These children did not have difficulty with auditory discrimination and distinguished readily between right and wrong pronunciation when spoken by the therapist. Auditory memory was normal for age when tested with nonsense syllables and the standard sentences from the Terman-Merrill Intelligence Scale. Drawing was good and they had no constructional or spatial difficulty.

The developmental history and clinical testing of motor maturation supported the diagnosis of delayed motor maturation. Walking was late. The mean age of walking of the group used in the investigation (see Method following) being 16·6 months, range 14-23, whereas the mean of the language group was 14·4 months, range 10–18.

Accurate dating of commencement of speech was not possible in every case, since the mother, whose report was the sole source of information, was not able to determine when pre-linguistic babbling merged into speech ; some mothers in both groups were sure their child had never babbled in the first year.

Clinical tests of motor maturation included closing fist and thumb, thumb and finger opposition, and cutting paper with scissors (Merrill Palmer) ; hopping on one foot (Gesell) and ability to protrude and lift tongue-tip (Morley 1957), besides ability to imitate articulatory movements in consonant and vowel syllables. All children exhibited signs of late motor maturation and poor co-ordination which appeared to stem from a difficulty in isolating movements, with delay in the normal synkinetic evolution as described by de Ajuriaguerra and Stambak (1955).

The Language Disorder Group.

The language disorder group did not suffer from a similar delay in motor maturation, they crawled, sat and walked within normal limits and performed well on co-ordination tests. They had little or no difficulty in direct imitation of consonant blends and words but had considerable difficulty in *recalling* the correct phonetic sequences and also grammatical constructions in spontaneous speech. Their speech was characterised by inconsistent substitutions and omissions which rendered speech often incomprehensible even to the mother. They made unpredictable mistakes in the order of sounds within words and words within sentences. Conjugation of verbs was poor and they peformed badly on the Watts language picture tests. They had difficulty in remembering names of things and people. They could discriminate by ear perfectly when pictures of similar sounding articles were named and when the correct pronunciation was contrasted with the child's error. They could not determine whether a word was right or wrong, if presented with their own erroneous version, as in the case of the child who said ' umblella ' or ' lumbella ' for umbrella.

Auditory memory as tested by repetition of groups of nonsense syllables and the standard sentences from the Terman-Merrill Intelligence Scale was deficient. Some of the children had a constructional or spatial difficulty and failed the Merrill-Palmer tests for copying a cross, a star and the three and six-cube pyramids and assembly of the Manikin. Drawing showed an inability to appreciate form and orientation. For these reasons it was felt that these children had a deficient sense of order in time and space associated with the difficulty in retention and recall of auditory material.

II.—METHOD.

The subjects were children with severe speech defects referred for treatment to the speech therapist (M.C.L.G.). None of them had any physical malformation; clinical and audiometric examinations by the appropriate authorities had revealed no neurological or auditory abnormality. Those who fell into one of the two categories described above were referred to the Psychologist (T.G.C.), who carried out the tests. The speech therapist made the differential diagnosis between motor and language before seeing the test results, but the diagnosis was not known to the psychologist until the whole series had been tested The children had to be at least five years old when the testing was carried out, because this is the minimum age for one of the tests used (the W.I.S.C.); also they were not sent until their speech was at least intelligible to a person unfamiliar with them. All appropriate cases were referred for testing until a total of 20 was reached (8 motor (7MIF) and 12 language (8M4F)). This took nearly two years. In the testing session, the following investigations were carried out:

W.I.S.C. The I.Q. was calculated from seven subtests: Vocabulary, Similarities, Arithmetic, Digit Span, Picture Completion, Block Designs and Coding. Any child with a full I.Q. below 85 was excluded, to avoid cases where speech disability was due to inadequate general intelligence. The subtests were used for the investigation of immediate memory and ordering ability. The advantage of having standardized scores was thought to outweigh the disadvantage of using tests which were not specially devised ' pure ' measures of the abilities concerned. From the way in which the scores are standardized, a child's Scaled Score on any subtests can be compared directly with his scores on other subtests, and also with the scores of other children of different ages. Digit Span was used as a measure of immediate memory, Arithmetic and Block Designs as measures of ordering ability, Arithmetic, at this level, being a measure of the child's familiarity with the sequence of numbers and the ability to apply it to simple addition, Block Designs measuring the ability to copy groups of objects in their correct spatial relationships. They were compared with Vocabulary and Similarities, which were considered the most free from ' ordering.' The other two subtests were omitted, as not falling definitely into either category. It was predicted that the Language group would be relatively poor on the Digit Span, Arithmetic and Block Designs, compared with their Vocabulary and Similarities.

Reading, etc. In Hallgren's (1950) large study of reading disability, the only factor of those studied which showed a highly significant relationship to dyslexia was the presence of speech defects. The reading of our cases was examined to see if it was poor, and to what extent. No hypothesis was made about the difference between the two groups. A mechanical reading test was used—Schonell's Graded Word Reading Test; this was thought preferable to a comprehension test, as being less dependent on intelligence. The reading age obtained from this was converted to a quotient in the usual way, so that it could be compared approximately with the I.Q. As it was not thought appropriate to test a child's reading before the age of 6, any child who was not yet 6 on the first testing was seen later, just for the reading, when he was at least 6.

Each child was asked to write his name and two simple sentences, ' a big dog ' and ' I like apples,' chosen to bring out reversals or transpositions of letters (what Orton (1937) calls ' static ' and ' kinetic ' reversals) since these are important to theories which link speech and reading disabilities with peculiarities of laterality. This was scored simply for the presence or absence of reversals of either kind. The letters were dictated to children who could not spell words.

Laterality. The history of the study of the relationship between speech and reading difficulties and various types of laterality (left handedness, mixed or crossed laterality, etc.), is a long and contradictory one, well summarized by Vernon (1957). No prediction was made as to differences in laterality between the two groups. The assessment was made by a simple group of observations of the ' choice ' type : hand used for writing, hand for throwing (five trials), hand for catching, when instructed to use one hand (three trials), eye for counting fingers looking through a small hole in a piece of paper held in two hands (three trials, in three directions), eye used in naming pictures through a ' V-scope ' (Miles' ABC Vision Test) (five trials), and foot for kicking (three trials). Clark (1957) found kicking the best and most consistent test of footedness ; a matchbox was used instead of a ball to facilitate central placement. There are twenty items in all, and they were scored just as ' consistent ' or ' inconsistent.' The criterion for ' consistent ' was nineteen or twenty items the same side, unless the one different was the hand used for writing ; anything less than nineteen was ' inconsistent,' though as it turned out the highest score in the ' inconsistent ' group was seventeen, and the ' consistent ' all scored twenty.

The children were also asked to indicate their left foot, right ear, etc. (five trials) and to perform crossed movements (e.g., " touch your right ear with your left hand ") (three trials). These two sections were simply scored as ' correct ,' if all trials were correct, or ' incorrect.'

III.—RESULTS.

Intelligence and Analysis of Subtests.

TABLE 1

INTELLIGENCE AND SUBTEST SCORES OF THE TWO GROUPS.

	Motor (N=8)			Language (N=12)				
	Mean	S.D.	Range	Mean	S.D.	Range	t	P
Age	7·2	1·55	5·3–11·1	7·1	1·33	5·1–9·8	—	—
I.Q.	113·8	8·57	105–131	101·2	10·50	88–116	—	—
Subtests .								
Vocabulary....	12·1	1·47	10–14	11·3	2·43	8–15	0·74	>·4
Similarities ..	11·3	2·22	9–15	10·7	2·21	7–14	0·54	>·5
Digit Span	11·8	2·11	9–14	8·7	1·49	6–12	3·63	<·01
Arithmetic	13·1	2·37	9–16	9·9	1·76	7–11	3·30	<·01
Block Designs	13·3	2·28	10–18	10·6	2·50	7–15	2·30	<·05

The two groups differ significantly in the expected direction on the tests where a difference was predicted, while differences in Vocabulary and Similarities are not significant. Unfortunately, the Motor group have a considerably higher I.Q. than the Language group, so that they would be expected to be significantly better on some subtests. To eliminate the effect of I.Q., scores on each of the three critical subtests were subtracted from the mean of Vocabulary and Similarities for each individual. If the two groups are compared with

TABLE 2

COMPARISON OF SUBTEST SCORES.

	$\frac{V.+S.}{2}$ — Dig. Span				$\frac{V.+S.}{2}$ — Arithmetic.				$\frac{V+S.}{2}$ — Blocks			
	M.	S.E.	t	P.	M.	S.E.	t	P	M.	S.E	t	P.
Motor	−0·06	·80	0·08	>·9	−1·44	·89	1·61	>·1	−1·86	·95	1·96	>·05
Language	+2·33	·58	4·00	<·01	+1·08	·61	1·77	>·1	+0·42	·64	0·65	>·5
t (comparing the two groups)	2·47				2·42				2·06			
P	<·05				<·05				>·05 (<·1)			
	(V.=Vocabulary ; S=Similarities.)											

respect to these differences, we have an estimate of their relative efficiency on the critical tasks, as compared with the neutral ones, I.Q. being held constant. Also, from the way in which the scaled scores are derived, we know that random groups of the population would be expected to have a mean difference of nought when any two subtests are compared ; so by examining the approximation of our group mean differences to nought, we can estimate whether they differ from the general population in their relative skill on these tests, I.Q. again being kept constant. Table 2 shows the result of these two operations.

It will be seen from Table 2 that the means of the Language group all deviate from nought in the predicted direction, but only one is significant, Digit Span. In comparing the two groups all differences are in the expected direction and two are significant at the 5 per cent. level. If a one-tail test is used (as is reasonable since the results were predicted) all three are significant at the 5 per cent. level. Thus, the Language group are poorer than the Motor group on the three tests predicted, but compared with the general population on only one test, Digit Span, is their low scoring significant.

Reading. Table 3 shows the mechanical reading quotients of the two groups compared with their I.Qs. One child was lost from each group ; these both did the original testing at the age of 5, and could not be obtained for a second assessment after their sixth birthday. Neither of these could read any words, and they only named about half of the letters in the first line of the test correctly (4 and 6 respectively out of 10, I.Qs. 105 and 116, ages 5 years 4 months and 5 years 6 months), so their reading was probably retarded. The Table shows that both groups were considerably backward at mechanical reading, and both to about the same extent. Four out of seven of the Motor group and five out of eleven of the Language had reading quotients more than 30 points below the I.Q. Comparison of the (I.Q.—R.Q.) means of the two groups gives t=·50, P> ·6.

TABLE 3

MECHANICAL READING QUOTIENTS.

	Motor Group (N=7)	Language Group (N=11)
Age	M=7·44 Range=6·3—11·1	M=7·64 Range=6·1—9·8
I.Q.	M=115·00 Range=105—131	M=99·82 Range=88—115
R.Q.	M=86·14 Range=72—107	M=73·64 Range=61—87
I.Q.—R.Q.	M=+28·86 Range=+13—+38	M=+26·18 Range=+8—+45

Laterality, Knowledge of Sides and Reversals.

All eight of the Motor Group were ' inconsistent ' on the laterality tests, but only four out of twelve of the Language Group. The probability of getting such an extreme division as this, by direct calculation (two-tailed) is ·008. There is no obvious pattern in the type of inconsistency shown by the Motor Group.

Putting the predominant choice for each member in the order hand, eye, foot, there are three RLR, two LLR, one RRL, one RLL and 1 LRL. Four of them also had inconsistencies within one member. In the Language Group there were eight RRR, one RRL, two RLR and one LRR. From the data given by Harris (1949) the proportion of inconsistent in the Language Group is about what would be expected in an unselected sample of children of this age group.

From Table 1, it can be seen that the Motor Group are older and more intelligent than the Language Group, so this lack of establishment of laterality cannot be ascribed to lesser general maturity.

The ability to name one's sides is probably an all-or-none ability. According to Launay (1952a), nearly all dyslexic children can do it by the age of 8 and nearly all normal children by the age of 7. Thus, it is unlikely to be discriminating after the age of 8, and we restricted this comparison to children between the ages of 5 years 0 months and 7 years 11 months. This left seven of the Motor and nine of the Language. Five Motor and only two Language were correct in simple indication of sides, and all seven were of inconsistent laterality. The nine who made mistakes were three inconsistent and six consistent. The probability of drawing all seven inconsistent from a group of ten inconsistent and six consistent (two-tailed) is ·021. Unfortunately, knowledge of sides in this age-group is clearly related to intelligence. The Mean I.Q. of the ' correct ' group is 112·3, of the ' incorrect ' group 98·8. However, there is certainly no evidence of a relationship between inconsistency of laterality and lack of knowledge of sides, and in any case the ' incorrect ' group is mostly drawn from the Language group whose I.Q. is thought to be lowered by their special disabilities. There were four 7-year-olds in the Language group, and they all made mistakes ; so they are probably worse at this than normal ; the three 7-year-olds in the motor group were all correct. Only three of the whole series were able to carry out crossed movements, two motor and one language, all three of inconsistent laterality.

In considering the question of reversals of writing, the second test of the original 5-year-olds was used. As two were missed, this gives six Motor and eight Language between the ages of 6 years 0 months and 7 years 11 months. Four of the Motor group and six of the Language showed reversals, five out of six with consistent laterality and five out of eight inconsistent. There is no relationship here between reversals and type of speech disorder or type of laterality. The degree of reversal varied greatly from just one error (' b ' for ' d '), to a complete confusion of all letters. This clearly needs a more systematic method of study.

IV.—DISCUSSION.

The prediction of the relatively poor performance of the Language group on Digit Span, Arithmetic and Block Designs was borne out in the case of Digit Span and Arithmetic. The Motor group were significantly better than the Language group on the Block Designs, but this was due to the fact that the Motor group were better on these than on other tests, not that the Language group were worse. Digit Span, the most discriminating test, is a test of ' immediate auditory memory,' but in memory tests it is always advisable to ask ' memory for what ?', and in this case it is memory for numbers *in order* ; mistakes are often mistakes of order, the correct digits being given but in the wrong sequence. So this can, perhaps, also be taken as evidence for the poor ordering ability of the Language group.

The most interesting finding is the difference between the two groups in consistency of laterality, the Motor group being all inconsistent, the Language group containing a ' normal ' number of children with inconsistent laterality. This seems to provide a further definite basis for distinguishing the two types, in addition to the speech differences. It may offer a lead to the solution of the problem of the contradictory findings about laterality in reading disabilities.

We have two groups of children, who can be characterized in the following way: Group 1—Speech disability of the ' motor ' type, poor reading, inconsistent laterality, no deficiency of immediate auditory memory or ordering ability; Group 2—Speech disability of the ' language ' type, poor reading, ' normal ' laterality, poor immediate auditory memory and poor ordering ability. In addition, knowledge of sides does not appear to be related to laterality, and tendency to make reversals in writing is not related to inconsistent laterality or to poor ordering ability as assessed by other tests; on these two points, however, the evidence is rather flimsy. In absolute terms, Group 2 are considerably worse readers than Group 1. Not one of them reaches a quotient of 90. The two groups show the same discrepancy between reading and I.Q., but it must be remembered that the I.Q. is partly composed of tests in which Group 2 are thought to do badly because of their language disability.

In spite of the negative evidence, there are so many studies showing a relationship between poor reading and peculiarities of laterality that there may be some truth in it. The easiest way of explaining the discrepancy in the findings would be to suppose the existence of two or more types of disability, and our Motor group appears to be one quite well-defined type, corresponding roughly to what Vernon (1957) calls " a class of individuals who are generally lacking in maturation " (page 109), though we should perhaps qualify it as ' motor maturation,' since they are of good intelligence, with little variation in abilities. The other group is not so well defined, the nature of their deficiency in ' ordering ' ability not having been demonstrated in a completely satisfactory way. Block Designs failed to bring out clearly the clinically observed spatial difficulties of the Language Group, although they were, in fact, significantly worse on them than the Motor group. This may be due to the fact that it is a copying test. It is well known that children with severe difficulty in writing letters from memory can often copy them quite well. Perhaps a creative rather than a copying type of test would be more discriminating; de Hirsch (1957) considers a draw-a-man test to be useful for this purpose. It is hoped to investigate this further with more refined procedures.

Much of the study of developmental language defects in children has been concerned with the association of three things: (*a*) the actual disability in speech, reading and writing; (*b*) peculiarities of laterality; (*c*) defects in spatial and temporal ordering. There are *a priori* reasons for linking (*b*) and (*c*) but it seems possible that it is the attempt to do so which accounts for some of the confusion; our results suggest two groups, one with (*a*) and (*b*), one with (*a*) and (*c*), with no connection between (*b*) and (*c*). The linkage of (*a*), (*b*) and (*c*) is usually associated with the notion of a ' dominant hemisphere '; the failure of one hemisphere to become completely dominant would result directly in (*b*), and also in (*a*), since it has long been considered that speech is somehow located in the dominant hemisphere; (*c*) could be thought of as arising directly from the lack of dominance, in the way suggested by Orton, or it could be secondary, a consequence of (*b*), which might make it more difficult for the child to develop a consistent sense of direction related to his own body. However, recent findings (Humphrey and Zangwill (1952), Penfield and Roberts (1959), Zangwill (1960)) suggest that the area concerned with speech tends to be in the left hemisphere in left-handed as well as right-handed people, in other words that it is not related to dominance. There is certainly a sense in which one hemisphere is (or becomes) dominant, if only as a description of the fact that most people are more skilful with one hand than the other, and tend to choose that hand (and the corresponding foot) for skilled operations. Our Motor group certainly lack dominance in this sense, so it looks as if motor skill in speech is related to speech defect ; after all, the written symbol is based on the spoken, so if the first symbol system is faulty, the second must suffer. The Language Group, on the other hand, could be considered as having a defect in a speech or language area, not related to dominance, or perhaps in some system concerned with a more general function. Their difficulties bear at least a superficial resemblance to Gerstmann's syndrome, their poor reading and writing, poor arithmetic and sense of order, and lack of knowledge of left and right. Drew (1956) considers both Gerstmann's syndrome and ' congenital word-blindness ' to be based on a failure in Gestalt formation. In contrast, the Motor group are good at Arithmetic, and their knowledge of sides good for their age (c.f. Launay (1952a)). It is an interesting feature of this study, and worthy of further investigation, that the expectation that confusion of laterality and poor knowledge of sides would go together was not confirmed. They appear to be unrelated.

Another source of confusion in this subject is the question of reversals in reading and writing. The evidence is convincing (Vernon (1957), Chap. V) that such reversals are related to poor reading but not to laterality. If a child is going to make a mistake in writing, it is quite likely that the result will be such that it can be interpreted as a reversal of some kind. It could be that attention is drawn to this kind of error to the exclusion of other kinds. This point is clearly brought out by Miles (1961), in his discussion of the mistakes made by two children with ' developmental aphasia.' It is clear from his analysis that the disorientation which he considers to be at the root of the problem is a highly complex thing, involving space and time. It is interesting to note that his two cases of severe disability were both of consistent laterality. It is well known that nearly all children show reversals at some stage of learning to read and write (c.f., Launay, 1952). It might be that some reversals are just mistakes, while others are symptomatic of a genuine disorder of orientation.

It is recognised that much of this discussion is speculative and goes beyond the facts. However, we feel that we have distinguished two types of developmental language disorder and that investigation with the two types in mind may help to clarify some of the issues.

V.—REFERENCES.

CLARK, M. M. (1957). *Left-handedness.* London · Univ. of London Press.

DE AJURIAGUERRA, J., and STAMBAK, M (1955). L'évolution des syncinésies chez l'enfant. *Presse Méd.*, 39, 817-819.

DE HIRSCH, K. (1957). Tests designed to discover potential reading difficulties at the six-year old level. *Am. J. Orthopsychiat*, 27, 566-576.

DREW, A. L. (1956). A neurological appraisal of familial congenital word-blindness. *Brain*, 79, 440-460.

GREENE, M. C. L (1960). Diagnosis and treatment of late speech and language in children. *Folia Phoniat*, 12, 101-107.

HALLGREN, B. (1950). Specific dyslexia. *Acta. Psychiat. Neurolog. Supp.*, 65.

HARRIS, A. J. (1947). *Harris Tests of Lateral Dominance.* New York : Psychological Corporation.

HUMPHREY, M. E., and ZANGWILL, O. L. (1952). Dysphasia in left-handed patients with unilateral brain lesions. *J. Neurol. Neurosurg. Psychiat.*, 15, 184-193.

AN ANALYSIS OF C. W. VALENTINE'S REASONING TEST FOR HIGHER LEVELS OF INTELLIGENCE

By H. J. HALLWORTH

(Department of Education, University of Birmingham)

SUMMARY. Answers to C. W. Valentine's reasoning test for higher levels of intelligence made by 286 men and 262 women were analysed and compared. Analysis of the item inter-correlations by the method of principal components gave a factor of general reasoning ability for both men and women. Upon rotation to the Varimax and Quartimax criteria three group factors were obtained which were comparable for the two groups of subjects. These factors are described and tentatively named. The results indicate the potential value of a detailed examination of the thinking processes involved in answering the items of this test.

I.—INTRODUCTION.

UPON the publication of his article on "The use of a new reasoning test for selection of university and training college students" (Valentine, 1961), Professor Valentine kindly made available for further analysis the test scripts of 548 university and training college students and sixth form pupils. This article reports the results of the analysis.

II.—METHOD OF ANALYSIS.

The scripts were divided into two groups, 286 completed by men, 262 by women. Using an electronic computer, two similar analyses were carried out, one for each group.

The test gives a score for each of sixteen different items. Including total test score and age, there were, therefore, scores on eighteen variables for each subject. Means, standard deviations and product-moment inter-correlations were obtained for all variables. The correlation matrix for the sixteen test items was analysed by the method of principal components, and the factor loadings were rotated to the Varimax (Kaiser, 1958) and Quartimax (Neuhaus and Wrigley, 1954) criteria.*

III.—MEANS AND STANDARD DEVIATIONS.

The mean age of the men was 22 years 3 months, and of the women 19 years 4 months. This difference was due to the fact that, at the time of testing, national service delayed the college entry of most men. The mean test score for men was 28·9, with a standard deviation of 14·5; the mean score for women was 24·1, with a standard deviation of 12·7. These figures are consistent with those given in the original article.

On fourteen of the test items men had a higher mean score than women. For two items women had a higher mean score, but in neither case was the difference between men's and women's means significant. It is, however, of interest that these two items are different from all others in that they each combine four characteristics:

(1) Each specifies one attribute of a logical class, using the words "None but . . ."

---

* All programmes used are for a Ferranti Mercury computer and were written by the author in mixed Autocode and machine code. (Hallworth, 1962.)

(2) Each specifies another attribute of the same logical class, using in effect the wording " All . . . "

(3) Neither of the two attributes is exclusive to the class.

(4) The answer demands an appreciation of the fact that the logical classes defined by each of the two attributes are not necessarily co-incident.

Precise information regarding adult sex differences in patterns of reasoning is not often available. This finding is recorded because, should comparable information appear from another sample using this test, or from another source, it would surely demand further investigation of the thinking processes involved.

As would be expected, the men had a higher standard deviation than the women for most items.

IV.—CORRELATIONS WITH TOTAL SCORE.

Correlations of total score with age were insignificant for both men and women, being —·05 and + ·09, respectively. The negative correlation for men was probably due to the inclusion in the sample of a number of sixth formers who had recently gained university open scholarships.

The correlation of each item with the total score is given in the table below. In general the figures are very satisfactory. The lower coefficients for the last two items are due to the greater difficulty of these items, a matter which is discussed again later.

TABLE 1

CORRELATIONS OF ITEMS WITH TOTAL SCORE.

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Men	·47	·47	·40	·50	·35	·43	·41	·39	·52	·54	·53	·58	·53	·41	·32	·31
Women ...	·47	·41	·41	·42	·43	·27	·30	·46	·43	·43	·48	·53	·50	·35	30	·13

V.—ANALYSIS OF THE ITEMS.

An obvious question about a test which measures reasoning ability at higher levels of intelligence is, does it measure one or more kinds of reasoning ability ? The generally accepted view is that reasoning ability depends largely on ' g ' (Valentine, 1961, p. 228), but that there may be differences between inductive and deductive reasoning, verbal reasoning and numerical reasoning, etc. In the original article, Professor Valentine indicates that the first four items of the test were constructed to measure inductive reasoning ; the remaining items are concerned with deduction. We should, therefore, predict a factor of general reasoning ability, and a factor or factors to differentiate between the first four and the remaining items.

(*a*) *A General Factor.*

From the correlation matrix of the sixteen test items for each of the two groups of subjects, eight principal components were extracted. These accounted for a total of 69 per cent. of the variance for the men's group and 67 per cent. for the women's group. In each case the first component could be regarded as a general factor, accounting for 22 per cent. of the variance for the men's group and 18 per cent. for the women's group. Item loadings on this component are given in Table 2.

TABLE 2

ITEM LOADINGS ON FIRST PRINCIPAL COMPONENT.

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Men	·49	·42	·39	·50	·28	·51	·31	·30	·69	·58	·49	·70	·69	·31	·22	·20
Women	·43	·44	·49	·40	·35	·29	·15	·41	·61	·42	·40	·69	·67	·27	·18	–·06

A first component of this kind is, of course, to be expected from this method of analysis. The question is, should the analysis be accepted as it stands, or should the axes be rotated ? Since there are always many possible sets of factors the decision to accept one set rather than any other must be made on psychological grounds. In the present case, acceptance of these components without rotation would give the general factor which has been postulated, and in this sense, the prediction is verified. In popular terms, this could perhaps be regarded as a typical British solution to the problem, extracting as large a common factor as possible.

However, the remaining components were not comparable across the two groups, nor did they make very good psychological sense in either case. It was, therefore, decided to rotate the axes in an attempt to discover another set of factors, more of which would have good psychological meaning.

(*b*) *Group Factors.*

For both groups, all eight components were rotated to the Varimax criterion. This is claimed by Kaiser (1958) to give results approximating to Thurstone's 'simple structure,' therefore, we may expect a more typically American solution. The results are given in Table 3.

Since the method of rotation tends to spread loadings more evenly between a number of group factors, the general factor has disappeared, and there are now three factors which are fairly comparable for both groups of subjects.

Varimax Factor 1 has its heaviest loadings on items 9, 12 and 13. In order to interpret this factor psychologically, we must first note that in section two of the test there are, as explained by Professor Valentine (1961), two kinds of item : first, items requiring ability to perceive a correct argument ; secondly, items requiring ability to perceive an incorrect argument ***and*** know ***why*** it is incorrect. Items 6, 9, 12 and 13 are the only items of the first type. It would probably be agreed that they can be answered more 'intuitively' than can the other times. The relatively low loading of item 6 may be accounted for by the greater complexity of argument in this item. Similarly, the relatively high loadings of items 10 and 11 may be accounted for by the brief wording of the argument and relative ease of answering. If so, the best general description of this factor is 'intuitive reasoning'.

TABLE 3

VARIMAX FACTORS FOR SIXTEEN TEST ITEMS.

	MEN				WOMEN		
FACTOR	I	IIA	IIB	III	I	II	III
Item 1	·29	·62	·15	−·12	·02	·69	·06
2	·11	·18	·76	·03	·11	·65	−·14
3	·07	·77	·05	·16	·36	·56	−·10
4	·30	·41	−·07	·10	·15	·35	·22
5	·01	·02	·14	·00	·05	·18	·13
6	·46	·08	−·22	−·07	21	07	·06
7	−·05	·07	·20	·10	·01	·01	−·08
8	·04	−·01	·13	·02	·16	·13	·14
9	·84	·06	−·01	·12	·77	·04	−·03
10	·53	·18	·23	·00	·20	·08	−·02
11	·29	−·08	·64	·08	·10	·23	·15
12	·66	·14	·19	·00	·74	·09	·00
13	·81	·04	·13	·03	·83	06	·07
14	12	−·06	.05	·11	·09	−·20	−·21
15	·11	−·12	·06	·80	·09	·11	·71
16	−·04	·27	−·01	·71	−·08	−·18	77
Square Sum	2·60	1·35	1·26	1·26	2·14	1·55	1·23

Varimax Factor II, for women, has heavy loadings on items 1-4, inclusive, and may be labelled 'inductive reasoning.' This factor was predicted. For men, the most comparable factor is IIA, which has, however, a low loading on item 2. Factor IIB is shown because this has a high loading on the missing item 2, also on item 11. There is no immediate logical similarity between items 2 and 11, but it may be remarked that on Factor II for the women, item 11 has the highest loading after the 'inductive' items 2-4. Here again, further investigation of the thinking processes involved might yield interesting results.

Varimax Factor III has its heaviest loadings on items 15 and 16 for both groups of subjects. These two items, the last of the test, are of a different order of difficulty from all others and yield a relatively low mean score for both men and women. Both involve relationships between five categories, and the answers involve double negation. Although Factor III was not predicted, it may be interpreted as relating to ability to deal with more complicated reasoning problems.

(*c*) *Further Group Factors.*

The Quartimax method of rotation has commonly been used for cognitive data on the assumption that it will bring out any common factor (Harman, 1960, p. 301). It was, therefore, used in the present instance in an attempt to obtain a general factor accompanied by other factors with psychological meaning. In Table 4 are listed the principal Quartimax factors for both groups of subjects. No general factor was obtained. Comparison with the Varimax factors shows a remarkable similarity. For both men and women, the sets of Varimax and Quartimax factors are almost precisely the same, differing by little more than the limits of precision written into the computer programmes.

TABLE 4

QUARTIMAX FACTORS FOR SIXTEEN TEST ITEMS.

	MEN				WOMEN		
FACTOR	I	IIA	IIB	III	I	II	III
Item 1	·29	·60	·16	—·16	—·01	·68	·04
2	·12	·17	·76	·02	·12	·65	—·14
3	·11	·78	·05	·14	·37	·56	—·12
4	·33	·42	—·09	·10	·18	·35	·21
5	·04	·04	·10	·00	·06	·14	·08
6	·48	·07	—·27	—·09	·19	·05	·04
7	—·02	·08	·15	·09	·00	·00	—·08
8	·04	—·10	·10	·02	·15	·15	·18
9	·84	·05	—·02	·11	·77	·02	—·04
10	·53	·17	·23	—·03	·21	·08	·00
11	·31	—·09	·61	·08	·10	·26	·17
12	·67	·12	·17	—·02	·73	·08	—·02
13	·82	·03	·12	·02	·84	·05	·05
14	·10	—·05	·04	·07	·09	—·20	—·21
15	·14	—·09	·06	·80	·10	·11	·68
16	—·04	·30	·00	·71	—·07	·16	·79
Square Sum....	2·69	1·34	1·20	1·24	2·15	1·53	1·29

VI.—CONCLUSION.

As is usual with factorial analyses, there is more than one set of factors which have psychological meaning. Accepting the principal components as they stand, there is a general factor of reasoning ability accounting for approximately one-fifth of all the variance. Rotating to either the Varimax or Quartimax criterion, there are three factors which are related to features of the test items and to the reasoning abilities which they apparently demand. These factors, and certain differences between scores for men and for women, suggest that it would be valuable to have a detailed examination, possibly along the lines used by Duncker (1945) and Peel (1962), of the thinking processes involved in answering the items of this test.

VII.—REFERENCES.

DUNCKER, K. (1945). On problem-solving. *Psychol. Monogr*, 58, No. 270.
HALLWORTH, H. J. (1962). Computer programmes for psychologists. *Bull. Brit. Psychol., Soc.*, 48, 16-21.
HARMAN, H. H. (1960). *Modern Factor Analysis*. Univ. Chicago Press
KAISER, H. F. (1958). The Varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.
NEUHAUS, J. O., and WRIGLEY, C. (1954). The Quartimax method. An analytic approach to orthogonal simple structure. *Brit. J. Stats. Psychol.*, 7, 81-91.
PEEL, E. A. (1962). Learning and meaning. *Bull. Brit. Psychol., Soc.*, 48, 1-9.
VALENTINE, C. W. (1954). Reasoning tests for higher levels of intelligence. Edinburgh : Oliver and Boyd.
VALENTINE, C. W. (1961). The use of a new reasoning test for selection of university and training college students *Brit. J. Educ. Psychol.*, 31, 227-231

THE DEVELOPMENT OF CHILDREN'S IDEAS ABOUT COUNTRY AND NATIONALITY

PART I: THE CONCEPTUAL FRAMEWORK.

BY GUSTAV JAHODA
(University of Glasgow)

SUMMARY. A random sample of children aged 6–11 in four Glasgow schools was interviewed and tested, two situated in middle-class and two in working-class districts. The formation of concepts relating the series of units Glasgow–Scotland–Britain is traced in terms of both verbal formulation and spatial representation. The difficulty for younger children of such concepts as 'town' and 'country' is demonstrated. The emergence of national identification is described, indicating the major sources of initial confusion.

I.—INTRODUCTION.

THIS study was originally inspired by Piaget (1928, 1951), who attempted to establish developmental stages for children's concepts of country and nationality. Unlike most other replications concerned with more clear-cut spheres of thought, such as the work of Lovell (1961) on mathematical and scientific concepts, the findings were in several fundamental aspects at variance with those of Piaget. The theoretical problems raised thereby will be dealt with separately. Here the aim is a two-fold one: to analyse the gradual emergence of concepts and relationships, and to offer tentative developmental norms for children of contrasting socio-economic background.

II.—METHODS.

(1) *The Sample.* The first stage of sampling was a 'purposive' one: two primary schools were selected from well-to-do middle class areas, where parental occupations are mainly clerical, professional and administrative; in contrast, two other schools were located in a high-density and now much dilapidated working class district, where semi- and unskilled occupations predominate.

Within the schools, a sampling frame was constructed from the registers, pupils being divided according to age and sex irrespective of their form. The minimum age was fixed at six, because it had been obvious from preliminary trials that it would have been difficult to administer the somewhat lengthy procedure to the youngest children; and it was felt that uniformity was essential if valid comparisons were to be made. After excluding any children not of Scottish parentage, selection was done at each age-level by means of a table of random numbers. three boys and three girls always forming the unit. This meant that the sample was widely scattered, only a small number being taken from any one form. The object of this was to minimize the possibility of communication among the participants, and there is every indication that this was satisfactorily achieved. The general plan is, perhaps, easier to follow when set out schematically:

Middle Class District (*M*) Working Class District (*W*)
School A School B School C School D
Boys Girls Boys Girls Boys Girls Boys Girls
(Ages: 6– 7– 8– 9– 10– 11–)

There were thus :

2 (social classes) × 2 (schools) × 2 (sexes) × 6 (age levels) × 3 (individuals) = 144.

Unless otherwise specified, all subsequent tables will be based on this total of 144 children. Although sex had to be one of the controlled variables, significant sex differences failed to emerge. Hence results will not be shown separately for boys and girls. This does not necessarily mean that there are no contrasts between the sexes in this field ; in fact, some of the differences were at least suggestive. However, in view of the relatively small numbers, dictated by the time-consuming nature of the interviews, it was felt that one would not be justified in putting much stresss on this aspect.

(2) *Interviewing and Testing.* The children were seen individually. Each session lasted between 35 and 70 minutes, the mean being about 45 minutes. In order to arouse and maintain the interest even of the younger children, the procedure was devised so as to present considerable variety.

In an investigation of this kind, the handling of the situation is of paramount importance. The procedure was not started until the children had been put at ease by some preliminary chatting, and, in fact, an atmosphere of informality was aimed at throughout. The schedule itself was indended as a guide, not a straight-jacket. The questions in the schedule were certainly always used in their exact wording ; but the possibility of purely verbal misunderstandings was constantly kept in mind. Hence, particular questions were often re-phrased and put again, taking care at the same time not to introduce any suggestions. The purpose was always that of exploring the degree of clarity and the range of various concepts for the child.

Naturally, this is a difficult task, and some influence of the observer upon what is being observed is unavoidable. As has been rightly said,* the very attempt to ascertain the child's capacity for comprehension is itself a learning situation. When a child is faced with a pointed question concerning some relationship for the first time, this may be the very stimulus which suddenly enables them to perceive how the elements fit together, so that a correct answer is produced after reflection. This did occasionally happen, but any error arising thereby will, from the present stand-point, be in the conservative direction. In other words, it will lead to an *overestimate* of the capacity to grasp a particular concept at any given age level.

III.—DEVELOPMENTAL FRAMEWORK.

(i) *Geographical Stages.*

There is general agreement that a child's intellectual grasp of his environment begins in his immediate vicinity and only gradually extends outwards. One can think of this as a series of concentric circles, with the child at the centre. Such a model, admittedly far too simple and neat to be a realistic portrayal of the child's perspective, may, nevertheless, serve to define the first task : it is that of determining the extent of each child's effective range. The word 'effective' means in this context the range within which the child is capable of correctly apprehending relationships.

One of the chief measures to be introduced here is what with some diffidence may be called the child's geographical stage, or 'G-stage' for short. This is derived from the verbally formulated relationships, in response to questioning.

---

* LOVELL and OGILVIE (1960).

The measure was arrived at after attempting, without successs, to apply Piaget's scheme. The one adopted here involves the use of what might be called ' natural ' units for the purpose of categorization. Whilst not eliminating the element of arbitrariness entailed in the slicing up and pigeon-holing of what is essentially a continuum, it does at least tend to reduce it. There were, of course, a number of marginal cases, but the large majority could be easily and unambiguously classified. The categories, or G-stages, were as follows :

G.I. : No conception of Glasgow as a unitary whole.
G.II : Conception of Glasgow as unitary whole, but no conception of it as part of Scotland.
G.III . Conception of Glasgow as part of Scotland, but no conception of Scotland as part of Britain.
G.IV : Glasgow-Scotland-Britain relationship correctly expressed.

The ideas of children at these four stages will now be illustrated with examples. The questions by the interviewer are shown in brackets ; the symbols after each quotation stand for type of school and age.

G.I.

It was characteristic of children at that stage that they seemed to think of Glasgow as some kind of vague entity near-by, though not usually including their own immediate geographical location.

(*Where is Glasgow ?*)
Beside the playground, in X street. (*W*6)
Its up by the park there—you go round the corner. (*W*6)
Down in Y avenue. (*M*7)
I don't know the way to Glasgow. (*W*7)

Although nearly all were familiar with the word ' Scotland,' their notions of this concept were apt to be nebulous in the extreme. Several main tendencies are distinguishable . (*a*) the fusion of Scotland with Glasgow ; (*b*) the belief that Scotland is somewhere *outside* Glasgow ; (*c*) random associations called up by the word ; and also combinations of these.

(*What is Scotland ? Where is Scotland ?*)
I get muddled up—is that not Glasgow ? Yes, it's just Glasgow. (*W*6)
My granny stays in Scotland. (*Do you go and see her ?*) Yes, I go up at 3 o'clock every day. (*W*7)
I was born in it—it's the name of the town.
It's just outside Glasgow. (*W*8)
Scotland is the capital of Edinburgh. It's in Glasgow. (*M*7)
Scotland is a street. (*Yes, Scotland Street, but have you heard of a place called Scotland ?*) Yes, Scotland the Brave, it's up in the Highlands. (*W*7)
The Kiltie-Dancers—on the telly—I saw it last night. (*Where is Scotland ?*) Away far. (*W*6)

The majority could say nothing about Britain at all, beyond perhaps the fact that they thought they had heard of it. If it was conceived as a place at all, it was clearly a very remote one.

(*What is Britain ? Have you heard of Britain ?*)
I've heard the word, but I don't know what it's like. (*W*6)
Yes. (*What do you think it is ?*) Our next-door neighbour went to it and brought some dolls home. (*M*7)

It was there that the men fought the war. (*W*8)

In general, it may be said that at this first stage the organization of geographical concepts is minimal, responses being mainly in the nature of diffuse and undirected associations.

G.II.

The chief difference between this stage and the preceding one is that the children's answers showed an awareness of being within the unit called Glasgow, which was no longer thought to be located somewhere else. Responses to the question " Where is Glasgow ? " were fairly stereotyped, usually taking the form " It's here, where we are." There were, moreover, a few replies which may be regarded as transitional to the subsequent stage, though they undoubtedly belong to G.II because of their inherent contradictions.

Glasgow is in Scotland. (*Where is Scotland* ?)
Scotland is a country outside Glasgow. (*W*7)
Glasgow is in the middle of Scotland. (*What is Scotland* ?) It's a place where the people play the bagpipes. (*Where* ?) Away in the Highlands. (*M*7)

The remainders' ideas about Scotland were largely similar to those of G.I, including the notions that Scotland is 'in England' or 'part of America' (both *M*6). A higher proportion ventured some statement about Britain, but the level of accuracy was usually not much higher.

(*What is Britain* ?)
It's a city in England. (*W*7)
It's a big town : mountains are there, and they climb on it. (*W*7)
I don't think I've been there. (*M*6)
It's a city in Scotland. (*M*9)

The major advance, therefore, is confined to the consolidation of the concept of Glasgow as an entity in which one lives and which constitutes one's surroundings.

G.III.

Here the stock answer became " Glasgow is in Scotland." or words to that effect. Scrutiny of the total response patterns suggests, however, that one must be careful not to read too much into this. It is true that replies did not go astray until after the question " Where is Scotland ? ", but as will shortly become apparent the foundation probably tends to be somewhat shaky.

(*What is Scotland ? Where is Scotland ? What is Britain* ?)
Scotland is a town made up of countries and towns. It's up north. Britain is a place where you live in, a big, big, piece of land surrounded by water. (*M*7)
Scotland is a country. It's in the world. Britain is really the whole continent. (*M*8)
Scotland is a country. It's beside Wales and Ireland. Britain is mostly called Great Britain ; it's sort of Scotland—another name for it. (*M*8)
Scotland is a country. Britain is the United State. (*What do you mean* ? Long pause.) Of America. (*W*9)
Scotland is a country. (*Where* ?) Can't say. Britain is a lot of different countries. (*What countries* ?) Glasgow, London, France. (*W*11)

When one considers these and similar answers the phrases " Glasgow is in Scotland " and " Scotland is a country " begin to have a suspicious air of an empty formula. Additional evidence shows that this is not always justified, yet in the main one seems to be faced with the parrot recital of terms learnt without adequate understanding. In other words, grasp of relationships has failed to keep pace with the acquisition of a new vocabulary. It is thus, perhaps, doubtful whether G.III should really always be accepted as a genuine stage regarding concept formation ; it might be safer to look on it as a temporary transition between G.II and G.IV.

G.IV.

This can be dealt with very briefly. In order to qualify, a child had to convey some understanding of Britain as well-defined unit. The first of the quotations below represents roughly the acceptable minimum.

> Glasgow is in Scotland. Scotland is a country in Great Britain. Britain is some countries joined together. (*M*6)
>
> Glasgow is in the Central Lowlands of Scotland. Scotland is a country. It's just north of England. Britain is Scotland, England, Wales and Northern Ireland. (*W*11)

With G.IV children have achieved the normal adult standard in Britain.

The levels reached by various sub-groups are summarized in Table 1. In this and most subsequent tables results are presented separately for three age-ranges, comprising twenty-four children each ; within age-ranges there is a further sub-division into *M* and *W* schools. This produces a general picture of developmental trends as influenced by age and socio-economic background.

TABLE 1

GEOGRAPHICAL STAGES REACHED BY TYPE OF SCHOOL AND AGE.

Geographical Stages	*M* Schools			*W* Schools		
	6–7	8–9	10–11	6–7	8–9	10–11
G.I.	4	1	—	16	5	—
G.II	13	3	—	8	10	3
G.III	4	7	2	—	4	12
G.IV	3	13	22	—	5	9

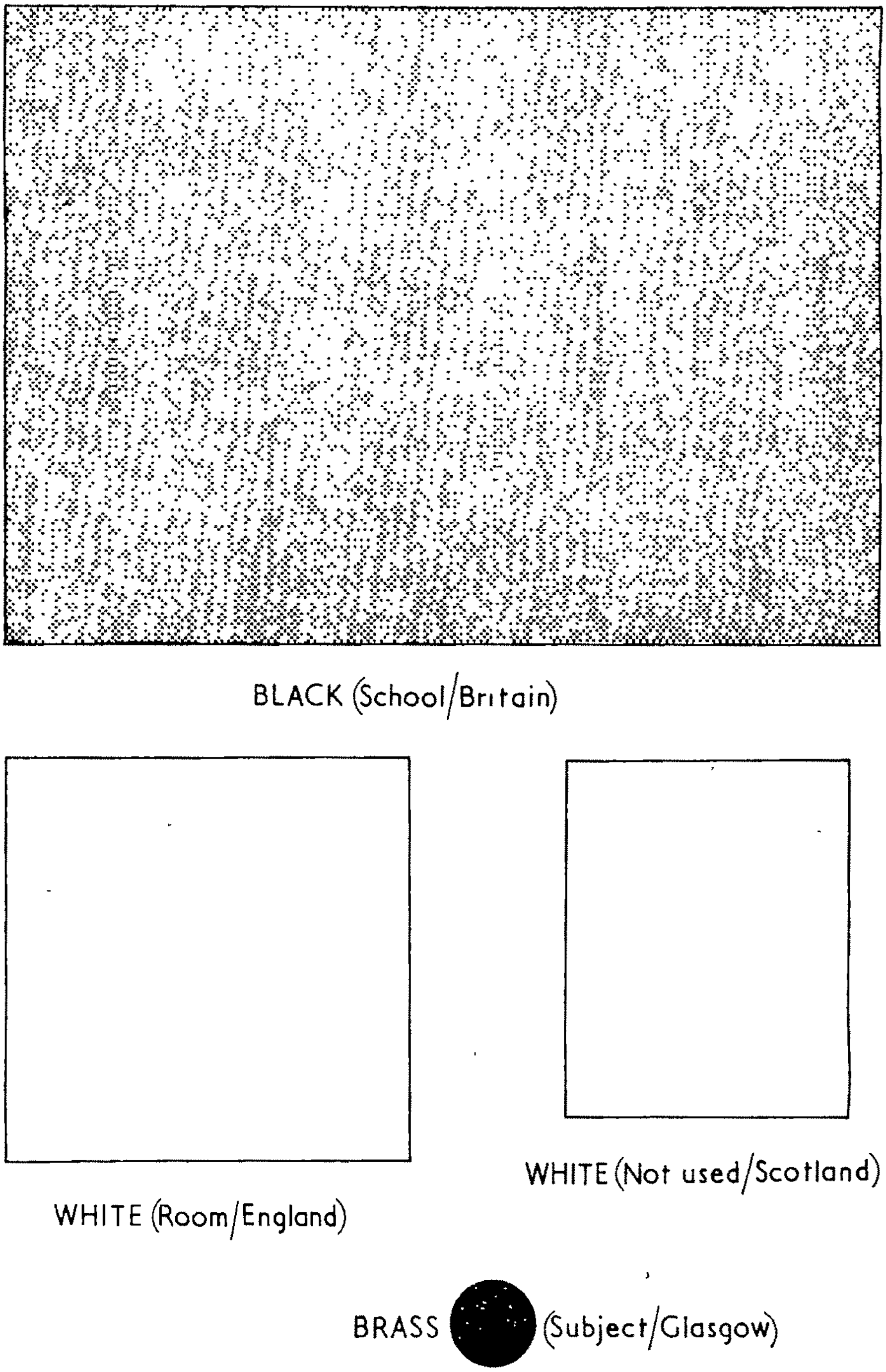
There is a highly significant difference (P less than ·001) between the distribution of stages in *M* and *W* schools. It may also be noted that G.III contains a smaller total than the stages preceding and following it, thus affording some support to the view that it might well be a relatively brief transition stage.

(ii) *Spatial Representation of Geographical Relationships.*

The procedure used in determining G-stages is open to the objection that it is conducted entirely at the verbal level, and in fact, some doubts have been expressed on this score as regards G.III. A special test dealing with the appropriate spatial concepts was, therefore, devised. It was administered in two parts the first being a preliminary training session designed to familiarize the children in a simple manner with the symbolic relations involved.

FIG. 1

Materials Used in the Space Test.



The material employed, seen in Fig. 1, consisted of one black and two white plastic boards, and a brass disc. For the try-out the smaller of the white pieces was left out, and the child asked to reproduce the immediate situation, i.e., self-in-room-in-School; this, is of course, analogous to Glasgow-in-Scotland-in-Britain. Any child failing in this preliminary task was given a detailed explanation and demonstration with the aid of the materials. This was necessary with 12·5 per cent. of all *M* children and 37·5 per cent. of *W* children (P less than ·001).

The decisive test came later, immediately following a musical test, so as to minimize carry-over from the questioning. This time the child was given all four pieces, with the information that the largest one was supposed to be Britain. The task was (*a*) to decide which of the others represented Glasgow, Scotland, and England, respectively, and (*b*) to assemble them correctly. The criterion for failure under (*a*) was simply not choosing the small brass disc for Glasgow; only a child making this mistake, whatever the subsequent performance, was classed as S.I.

In order to qualify for success in positioning, three conditions had to be satisfied: brass disc on top of one of the white pieces; white pieces side by side, not on top of one another; white pieces within the area of the black one. Anything not conforming to this was again classed as failure, S.I. The most common failures are illustrated in Fig. 2. Among the youngest children, side-by-side placement of all the pieces often happened (S.I.—A); overlapping sometimes occurred (S.I.—B), or simple building up (S.I.—C), which was an attempt to reproduce the earlier practice trial whilst coping with the additional piece. The most frequent failure of all was something like S.I.—D, namely, the brass disc flanked by the two white pieces; from the conceptual point of view, this suggests that such children thought of Glasgow and Scotland as two separate units, without as yet any notion of spatial inclusion.

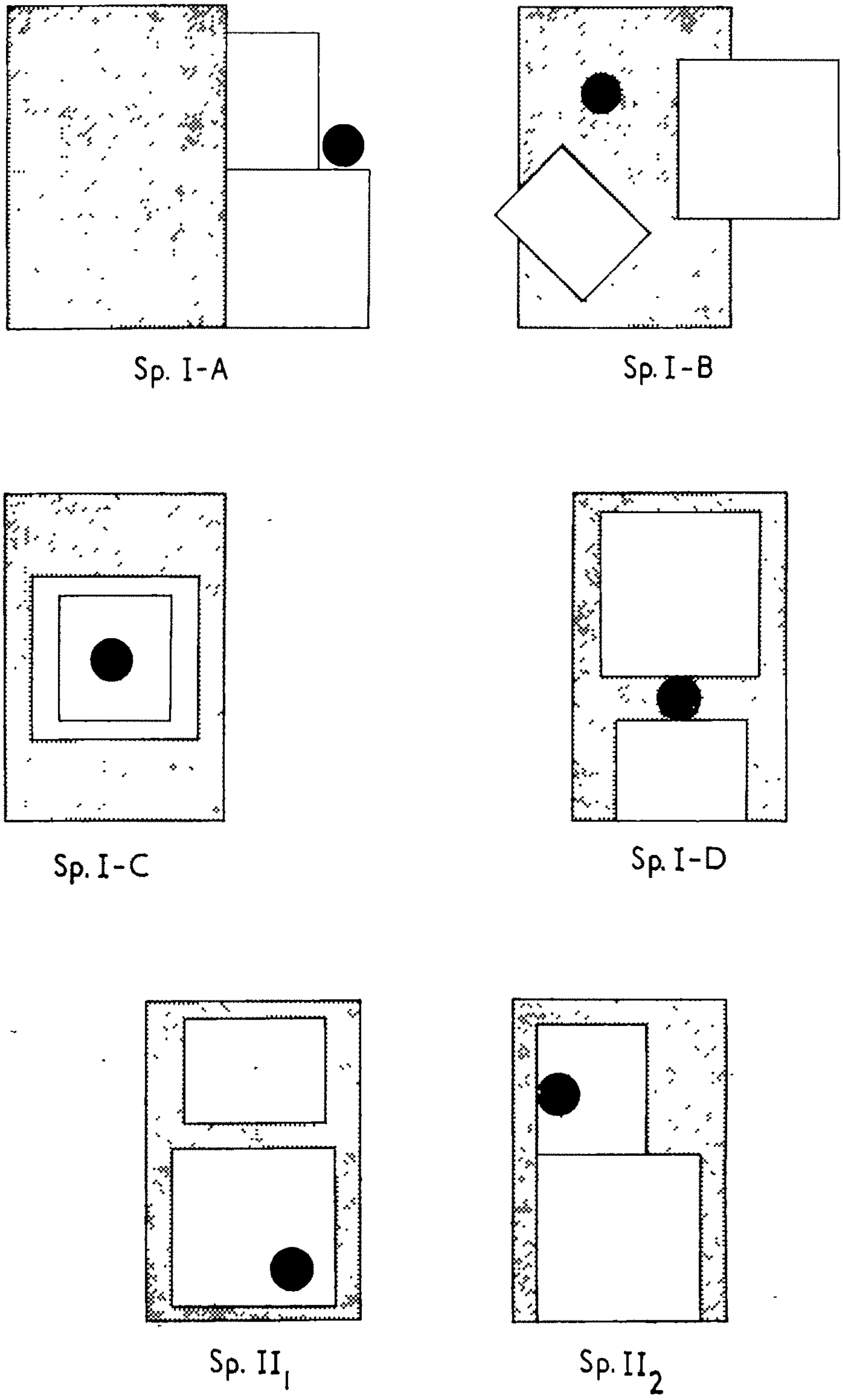
All children fulfilling the above conditions were scored as successes (S.II), ignoring both identification of white pieces and compass orientation. However, a further minor distinction was made between children who had at least one of these wrong ($S.II_1$), and those who were right in both ($S.II_2$). In the example of a successful solution shown in Fig. 2, the location of the Glasgow disc within Scotland was obviously carefully chosen; but the exact positioning was not considered in the scoring, as it would have been difficult to know where to draw the line.

Results are set out in Table 2, and it appears that the contrast between *M* and *W* is as sharp as it was in the case of G-stages, the overall difference being highly significant (P less than ·001). Closer scrutiny reveals a distinct age pattern: at 6–7 there is little to choose between *M* and *W*; but whilst *M* children progressed steadily from then onwards, *W* ones exhibited no appreciable advance until after the age of 9, and then they were still far from having caught up.

Another observation may be made concerning the relative frequencies of $S.II_1$ and $S.II_2$: it looks as though the former were less stable. In other words, once the basic nature of the Glasgow-Scotland relationship is grasped, compass orientation and size soon come to be sorted out as well.

FIG. 2

TYPES OF FAILURE AND SUCCESS IN THE SPACE TEST



Sp. I-A

Sp. I-B

Sp. I-C

Sp. I-D

Sp. II_1

Sp. II_2

TABLE 2

SPATIAL STAGE REACHED BY TYPE OF SCHOOL AND AGE.

	M Schools			*W* Schools		
Spatial Stages	6–7	8–9	10–11	6–7	8–9	10–11
S.I	20	8	4	22	21	12
S.II$_1$	1	6	4	2	2	3
S.II$_2$	3	10	16	—	1	9

(iii) *Comparison of the Two Approaches.*

This is provided in Table 3, which shows a regular and orderly progression. The relationship is most clearly apparent from the bottom row, which gives the percentages of complete failure on the space test for each G.-stage. The first sizeable reduction in the failure rate occurs at G.III, and there is a spectacular drop between G.III and G.IV.

TABLE 3

RELATIONSHIP BETWEEN GEOGRAPHICAL AND SPATIAL STAGES.

S-Stages	G-Stages				Totals
	G.I	G.II	G.III	G.IV	
S I	25	35	21	6	87
S.II$_1$	1	1	5	11	18
S.II$_2$	—	1	3	35	39
Totals	26	37	29	52	—
Per cent. S I. for each G-stage	96%	95%	72%	12%	—

The comparison offers a satisfactory confirmation of the interpretation already put forward. Stage G.III is clearly one of transition, in the course of which the verbal statement " Glasgow is in Scotland " remains largely an empty formula, devoid of an adequate conceptual appreciation of the part-whole relationship involved. On the other hand, a large majority of those able to given expression to the broader Glasgow-Scotland-Britain pattern have, in fact, achieved such a grasp.

Before leaving this important section, attention should also be drawn to the wide variations in ages at which a particular stage is attained, which emerges from Table 1 and 2. It is noteworthy, moreover, that only half the *W* children of 10 and 11 have reached an adequate concept of spatial relations, a fact which tends to be concealed by the effects of rote learning.

(iv) *The Concept of ' Large Town.'*

It may have been observed in connexion with some of the quotations that children were apt to describe countries such as Scotland and England as 'towns.' The converse error was also made occasionally, though there were fewer opportunities for this; a confusion of the kind is certainly implicit in such statements as "Scotland is the capital of Edinburgh."

Immediately following the topic of Glasgow, children were asked to name some other 'really big towns.' This was an attempt to discover the range of notions conveyed by the phrase at different ages. In evaluating the findings, summarized in Table 4, one limitation of this approach must be kept in mind: the total number of responses given by each child was not controlled, and this is a possible source of variation. However, as the general pattern seems to make good sense, it cannot have been of great importance.

Among the youngest children, nearly one-third were unable to give any answer, or mentioned street names; the remainder of this group produced a large and ill-assorted variety of responses, only two of them entirely correct. In the intermediate age-group, most of the crudest errors had dropped out, leaving mainly a confusion between small towns, large towns and countries. Finally, at the ages 10-11, the chief remaining mix-up was between small and large towns. As before, *M* children were generally greatly in advance over *W* ones (P less than ·001).

TABLE 4

TYPES OF RESPONSES FOLLOWING REQUEST TO NAME 'REALLY BIG TOWNS.'

	M Schools			*W* Schools			Rounded Percentages
	6–7	8–9	10–11	6–7	8–9	10–11	
Large towns only*	2	4	11	—	2	8	19%
Large and small towns	5	9	10	—	3	3	21%
Countries, large towns, and small towns	8	7	3	5	8	9	28%
Countries, small towns and streets	5	4	—	8	8	3	24%
Streets only	—	—	—	6	—	—	
Unable to name anything	4	—	—	5	3	1	9%

On the basis of these findings, one may hazard some inferences about elements of the thought processes involved. With the least mature group, the question evokes just any kind of geographical names that happen to be familiar to the child; in other words, the concept of large towns is so diffuse that it is taken to cover a variety of units, from streets to countries. If these concepts have not yet become differentiated, one can hardly expect the child to establish any kind of relationship between them, as was the task with regard to Glasgow, Scotland and Britain.

Gradually some boundaries begin to take shape, and the child moves into a zone of transition. Whilst the difference between towns and countries is not fully understood, some children are able to produce at least an appropriate verbal formula to designate the relationship (e.g., "Glasgow is in Scotland") others, though fewer, are beginning to apprehend the spatial relations.

---

* A 'large' town was defined as having a minimum population of 150,000.

The final stage, in terms of the classification adopted here, is the achievement of a sharp discrimination of units. This permits both an insightful verbal formulation, as opposed to parrot recital, and correct abstract representation of spatial arrangements.

If such an interpretation is valid, it follows that the notion of what constitutes a ' really big town ' ought to be closely associated with both G and S stages. Table 5 indicates that this is, in fact, the case (P less than ·001).

TABLE 5

TYPES OF ' REALLY BIG TOWNS ' RESPONSES IN RELATION TO G AND S STAGES.

	Response Categories		
	Large towns only + small and large towns	Small towns and/or large towns and countries	All other responses
Numbers in each category	57	46	41
Per cent. in G III and G IV	81%	54%	24%
Per cent. in S II	67%	37%	5%

IV.—COUNTRY AND NATIONALITY.

(i) *Spontaneous National Identification.*

The first question asked when the interview proper started was " What are you ? " The object was to elicit the dominant manner in which children tended to view themselves, before anything else could introduce a bias ; in particular to find out to what extent, if any, nationality entered into what are commonly called children's ' self-images.'

Results can be put very briefly : some nine-tenths confined themselves to their sex-role, i.e., they said " I am a boy/girl." Further probing brought out mainly qualifications of the statement already made (e.g., " I am a schoolboy," " I am a good girl "), or with older and/or brighter children, some superordinate categorization like ' person ' or ' human being.' Out of the total sample only just over 3 per cent. described themselves as ' Scottish,' all being from *M* schools. It may also be mentioned that the same proportion referred to their religion.

This may be compared with the findings of a study by Lambert and Klineberg (1959) carried out about the same time, and using the identical question. Three sets of English children (among other nationalities) aged 6, 10 and 14, respectively, were included in the survey. Omitting the oldest, the combined proportion of the others who called themselves English came to about 3 per cent., almost exactly the same. This is, perhaps, somewhat surprising, as one might have anticipated that Scottish children would show a relatively greater awareness of their nationality. It is possible that such a difference would appear later, as the 14-year-old English children showed a great upsurge in the frequency of giving their nationality.

(ii) *General Inquiry about Country and Nationality.*

For most adults the phrase ' my country ' has powerful emotional connotations. The problem of the growth of such sentiments is merely touched upon in this investigation, which focusses primarily on cognitive development. It was concerned, however, with the emergence of understanding of the object to which such sentiments refer.

Before Scotland or Britain had been mentioned by the investigator, children were asked " What is your country ? " It is apparent from Table 6 that the youngest children often had no idea, some naming either their street or local district. A frequent answer was ' Glasgow,' and this may have been influenced by the preceding talk about the home town. With rising age, ' Scotland ' gained increasing prominence, whilst the small proportion who named Britain remained fairly constant. Differences between *M* and *W* schools were significant (P less than ·01). The general pattern fits in well with the view that children begin with an awareness of their immediate locality, which becomes progressively extended outwards, *pari passu* with a sharpened discrimination concerning such geographical terms as ' town ' and ' country '

TABLE 6

DISTRIBUTION OF ANSWERS TO THE QUESTION ' WHAT IS YOUR COUNTRY ? ' BY TYPE OF SCHOOL AND AGE.

	M Schools			*W* Schools		
	6–7	8–9	10–11	6–7	8–9	10–11
Don't know or miscell	4	2	—	12	7	3
Glasgow	7	2	—	3	4	4
Scotland	11	19	22	8	11	15
Britain	2	1	2	1	2	2

The question " What is your nationality ? " held few surprises. Among *M* children only one in the 6-7 group understood it, as compared with one-third of those aged 8-9 and three-quarters of the oldest ones ; corresponding figures for *W* children were, respectively, nil, a single one and one-third. Most of those who understood the question gave a correct reply, i.e., ' Scottish ' or ' British,' but a few said ' England ' or ' English.'* The probable reason for such answers will be encountered shortly.

(iii) *Understanding of Relationships.*

After reference had been made to Scotland and Britain, children were asked specifically whether they were Scottish, and whether they were British. As is evident from Table 7, a majority of all ages gave an affirmative reply to at least one of these questions. Moreover, there is no significant difference between *M* and *W*, indicating that a verbal knowledge of nationality is acquired at roughly the same rate in both kinds of environment. Most of those aged 10-11 acknowledged that they were both Scottish and British ; the problem is, however, to what extent this corresponded to a grasp of the implicit part-whole relationship.

---

* It will be remembered that children not of Scottish parentage had been excluded from the sample.

TABLE 7

DISTRIBUTION OF COMBINED ANSWERS TO THE QUESTIONS 'ARE YOU SCOTTISH? ARE YOU BRITISH?' BY TYPE OF SCHOOL AND AGE.

	M Schools			*W* Schools		
	6–7	8–9	10–11	6–7	8–9	10–11
'No' or 'Don't know' to both	6	2	—	9	7	—
Scottish 'Yes' and British 'No' or 'Don't know'	8	4	—	5	4	4
British 'Yes' and Scottish 'No' or 'Don't know'	2	1	—	3	3	1
'Yes' to both	8	17	24	7	10	19

In order to test this, a follow-up was used with children who had said they were both, and with those who were positive about one but uncertain about the other. All these were asked about the possibility of being Scottish and British at the same time, and invited to explain their reasons.

The outcome, summarized in Table 8, is instructive. The most frequent reason offered for rejecting the possibility was mutual exclusiveness—" you can't be two different things at once." The other prominent feature, used as a justification on either side, was language. In fact, language and nationality were commonly mixed up, which probably led some of the children to say 'English' when asked about their nationality, as mentioned earlier. A few quotations will illustrate the prevalent state of confusion, the first pair below having answered 'Yes' and the second 'No':

> British is your language and Scottish is your accent. (*M*8)
>
> Down in England you speak British and here in Scotland you still talk British. (*W*10)
>
> I speak English—I'm really Scottish. (*M*6)
>
> You can't talk Scottish and British at the same time. (*W*9)

The first two happen to be consistent, though obviously without any glimmer of understanding; the last two, having declared previously that they were Scottish and British, had changed their position, yet there was no indication that they were aware of it. On the other hand, a small number of children did show such an awareness. For instance, one decided " You can't talk different ways," and reverted to Scottish only; another, possibly with a notion of Gaelic at the back of his mind, came to the conclusion that he was British but not Scottish because " All the Scottish speak a different language to us "; one altered his answer from 'No' to 'Yes,' explaining " Well—you can speak all the different languages." Lastly, an 11-year-old, after an initial emphatic 'No,' hesitated and finally reached the correct answer. This is an example of the interview as a learning situation, in the course of which new insights may be gained.

TABLE 8

IDEAS ABOUT THE RELATIONSHIP BETWEEN 'SCOTTISH' AND 'BRITISH.'

Can one be both ?	Explanation given	N
Don't know	None	4
No	None	7
	'Can't be two things at same time'	16
	Different language	6
	Other	3
Yes	None	4
	Same language	7
	Other	3
	Correct explanation	44

The miscellaneous answers contained some curious specimens. Some thought that Scottish and British were just different names for the same country ; one explained that his father was Scottish and his mother British, and a 6-year-old hit upon an ingenious solution : " One week I'm Scottish and the next week I'm British."

The lesson of this follow-up is that replies to the nationality questions cannot be taken at their face value. When challenged, about a third of the children said that one could not be British and Scottish at the same time, thereby mostly contradicting their earlier responses and showing that these had not been rooted in a genuine understanding ; others continued to affirm the possibility of being both, yet without being able to provide an adequate justification ; and less than half were able to demonstrate their grasp of the underlying relationship.

V.—REFERENCES.

LAMBERT, W. E., and KLINEBERG, O. (1959) A pilot study of the origins and development of national stereotypes. *Internat. Soc. Sci. J.*, 11, 221-238.

LOVELL, K (1961) *The Growth of Basic Mathematical and Scientific Concepts in Children*. London · University of London Press.

LOVELL, K., and OGILVIE, E. (1960) A study of the conservation of substance in the junior school child. *Brit. J. Educ. Psychol*, 30, 109-118.

PIAGET, J. (1928). *Judgment and Reasoning in the Child*. London . Kegan Paul

PIAGET, J., and WEIL A. (1951). The development in children of the idea of the homeland and of relations with other countries *Internat. Soc. Sci. Bull.*, 3, 561-578.

(*Manuscript received,*

PSYCHO-LOGIC: A POSSIBLE ALTERNATIVE TO PIAGET'S FORMULATION

By G. HARRY McLAUGHLIN
(Northampton College of Advanced Technology, London)

SUMMARY. The limitations of children's reasoning observed by Piaget can be described by a restricted logical system, called a psycho-logic. It is shown that psycho-logic can be simply represented by a small set of finite Boolean algebras. These are interpreted to suggest that limitations of reasoning correspond to limitations of immediate memory or attention.

PIAGET (1947) is particularly noted for the discovery that children behave as if their thinking conformed to the rules of a psycho-logic, meaning a logical system too restricted to permit certain valid deductions to be made. He also observed that a child's ability to reason progresses through an invariable sequence of four main stages, outlined below. He accounts for this sequence by suggesting that the development of higher stages of reasoning depends upon the child achieving concepts which are more and more abstract.

Nearly all the generalisations on the development of reasoning given by Piaget (1937), Piaget and Inhelder (1941, 1959) and Inhelder and Piaget (1955), have been supported by numerous careful investigations, such as those frequently reported in this *Journal* by Peel and his co-workers.

To describe the four stages of intellectual development, Piaget formulates four levels of psycho-logic. Each level is a logic so restricted that it yields only those inferences which can be made by a subject at the corresponding stage. Furthermore, Piaget advances the hypothesis that if, at any level, a formula can be derived from another formula in psycho-logic, a subject at the end of the appropriate stage can carry out cognitive processes corresponding to both formulas.

Psycho-logic has been set out in detail by Piaget (1949); particular aspects of it have been exhaustively treated by Piaget (1942, 1952); and it is summarised by Piaget (1947, 1953). The last-mentioned volume has an introduction by Mays explaining with admirable clarity the logical terms used in psycho-logic. Unfortunately, as Beth (1950) noted, Piaget's own expositions of his system contain serious confusions and are consequently hard to follow. Grize (1960) has expressed the system in a rigorous form, but this is even more formidable to the non-logician, and few of its terms appear to have psychological referents.

The difficulty of Piaget's system seems to stem from his attempt to translate directly into logical terms the psychological notion of increasing conceptual abstractness. This gives the system a structure different from that of conventional logics.

However, it will be shown that a psycho-logic can be adequately formulated in terms of the algebra of classes, which is explained in most introductions to modern logic (e.g., Langer 1953). It will also be shown that each level can be simply defined by the number of different classes that are distinguished simultaneously. Furthermore, this number can be directly identified with a clearly measurable psychological characteristic—memory span.

Briefly, the argument is that a subject solving an intellectual problem must retain a number of items at the same time, the number depending on the logical complexity of the problem ; and that the number of items a child can retain simultaneously increases as he grows older ; therefore, as the child matures, he can solve more complex problems. In other words, each of the four stages in the development of thought is characterised in part by the number of items that a child can retain at the same time. It should be noted that increases in retentivity are claimed to lead to changes only in the logical aspect of thinking : no doubt changes in the quality of thinking in its other aspects involve further factors. Furthermore, an ability to retain a certain number of items is claimed only to be a necessary condition for making certain kinds of inference : it is not proposed as a sufficient condition.

There are at least two ways in which the number of items retained simultaneously could increase with age. The older child might either be able to remember more items after a single presentation, or be able to attend to more items while they are being presented. Even more probably, memory capacity and attention capacity increase together, perhaps because they are two aspects of the same process. For convenience of exposition, a child's growing ability to consider a certain number of items simultaneously will usually be treated simply as a function of enlarging immediate memory capacity, but it must be emphasised that this notion is not fundamental to psycho-logic.

The size of memory span is greatly affected by redundancy of the stimulus material. That is to say, if an item in a list to be recalled is deducible from other items, a subject's memory span will be increased by that redundant item. However, we will be concerned with items that are by definition non-redundant : these G. A. Miller (1956) has named chunks. Each digit in a digit span test forms a chunk, unless the subject is capable of mentally combining the digits into groups. Thus, from the norms of digit span tests (e.g., Wechsler 1949, 1955) we can say that the immediate memory can simultaneously retain two chunks at about 2½ years of age, four by 4½, and seven by 13½, falling to six at about 55 years of age. Thus, immediate memory capacity increases and declines in the same way as intelligence, correlating between ·42 and ·62 with full Wechsler Intelligence Scale scores. The reason why the correlation is no higher may be that, although a given stage cannot arise until the capacity to retain enough chunks has developed, it is not consistently manifested until much later, when the child has learnt to combine and separate the chunks mentally.

Let us call that part of his environment which a subject is considering at any moment his universe. This universe can be characterised by the algebra of classes. In this logical system, any collection of entities, physical or mental, is called a class. A class can consist of a single entity. A class with no members (e.g., the class of unicorns) is said to be empty. There can also be classes formed of classes. If A and B are classes, their logical product is the class of entities in both A and B ; and their logical sum is the class of entities in A or B or both. Classes which are parts of a larger class are called its subclasses. The universe class comprises all the classes relevant to a given proposition. The complement A′ of a class A consists of all entities in the universe class that are not in A.

We may now define a subject's concept of a class as the specification by which he decides whether or not an entity is in that class.* To put it another way, a subject's concept of a class consists in his knowledge of its distinguishing attributes. Gestalt psychology shows that we can form a single unitary concept of an entity or class of entities with any number of attributes. However, the subject, by selectively attending to any combination of attributes, forms as many concepts as there are combinations to consider. For example, if he sees a big white swan, he may apprehend the object either under the concepts ' swan,' ' big (thing),' ' white (thing),' ' big white (thing),' ' big swan,' ' white swan ' or ' big white swan.' Furthermore, the subject might restructure the concepts ' swan ' and ' black ' to form a concept of their product, which would be manifested by some such remark as : " Are there black swans ? " He might also restructure the concepts ' swan ' and ' raven ' to form a concept of their sum, for which he would probably know the name ' bird.' By selective attention to aspects of the concept of a sum or of a product, the subject could again form the component concepts ' swan,' ' black,' ' big,' etc.

---

* Again for simplicity of exposition, it is assumed that a subject categorises a given entity with certainty as either being or not being in a given class. If, however, he categorises an entity as possibly being in a given class, the argument of this paper still stands, but the logical apparatus required is the modal calculus noted by Lewis and Langford (1932, p 501) as being equivalent to that discussed here.

Now, making a deduction reduces to solving a problem of the type : given that certain classes are or are not empty, what other classes are or are not empty ? At the worst, all the classes defined by the attributes of the given classes must be considered. The maximum number of classes that might need to be considered is easily calculated. Suppose the given classes are distinguished by a number, N, of different attributes (or sets of attributes that occur together), then each relevant, distinct class is defined by one of the 2^N possible combinations of the presence or absence of those attributes. For example, the classes given by the premisses of a syllogism vary in three attributes, so finding a conclusion might involve considering up to $2^3=8$ classes. In practice, many problems are solvable without considering all the classes that might be relevant.

As Bernstein (1935) has shown, the importance of N is that for each value it is given, it specifies a unique logic, called a finite Boolean algebra. We will, therefore, label each level of psycho-logic by the appropriate value of N—which must be a positive integer, unless it is 0, as will be explained below.

Because the 2^N classes defined by N attributes are non-redundant, the concept of each class corresponds to a chunk. Therefore, the level at which a subject could reason can be inferred from his digit span, always presuming that this is equivalent to his memory span for concepts : the subject will be at level 1 if the span is 2 ; at level 2 if it is 3 or 4 ; and at level 3 if it is from 5 to 8. As already mentioned, finite Boolean algebras would still be valuable as descriptions of the logical status of reasoning at the various psychological stages, even if stages and memory spans did not correspond. But if there were no correspondence, the hypothesis that a certain memory span is a pre-requisite for reasoning at a certain level would have to be discarded. However, a search of the literature, including most of that cited by Munn (1954, p. 412 ff.), has failed to reveal evidence of children making inferences while not having the memory spans required by the hypothesis. Indeed, remembering that memory capacity is only a necessary condition for reasoning at a certain level, and that success cannot be assured without the maximum capacity required at that level, it will be shown that there is a satisfactory agreement between the ages at which the various memory spans are attained and those at which the corresponding levels begin.

The following sketch of levels and stages uses those of Piaget's hypothetical constructs that are clearly supported by empirical data ; terms for these constructs are italicised.

LEVEL 0. This demands the capacity to process $2^0=1$ concept at a time. It corresponds to the *stage of sensorimotor intelligence*. This is defined by Piaget (1937) as the period up to two years of age, during which a baby learns to form *object concepts*, that is, to treat a given object as the same object however, and whenever, it is perceived.

We may suppose that some conditioning process leads the baby to categorise identically percepts similar in important respects. It is not known exactly how object concepts are attained. Piaget (1957, pp. 86-9) has suggested that a baby twiddling some object around finds that there is an invariant in the group of spatial displacements he has performed upon the object, and thus learns that the object is permanent: yet it must be pointed out that any finite number of twiddles may not form a group, and how invariants are extracted from perception remains unexplained.

For present purposes, it suffices to remark that the baby can make no inferences. This is due to the fact that an inference, by definition, involves relating at least two concepts, and at level 0 only one concept can be differentiated at a time. It may now seem paradoxical to suggest that the baby's object concept has $N=0$ distinguishing attributes. The paradox vanishes, however, when it is noted that if he has only one concept in mind there is no other concept to distinguish it from, therefore, distinguishing attributes are meaningless in this context.

LEVEL 1. This level demands the capacity to retain $2^1=2$ concepts simultaneously. It corresponds to the *stage of pre-operational intelligence*. This lasts from two to seven years of age, and is defined by Piaget and Inhelder (1941) as the time during which the infant comes to form *class concepts*, that is, to classify objects.

At first the infant may fail to note the distinguishing attributes specified in an adult's concept of a given class, so he may make classifications which adults think mistaken. To overcome this he will, presumably, try to learn the distinguishing attributes of all manner of classes; this he seems to be doing when he asks a series of questions of the form: "What's that? A cow? Why? What's it for?"

A more serious source of confusion is the infant's inability to distinguish between more than an object and its environment. He cannot at the same time note the relationship between the object and its environment, because that would necessitate the simultaneous retention of a third concept. But knowing the relative position of an object is essential to deciding whether percepts of similar objects encountered at different times should be classified as a number of objects of the same class, or as the same object. Piaget (1937, p. 374) has observed that infants are often in doubt on this score.

On the other hand, the infant's capacity to retain two concepts at once enables him to draw comparisons, which in turn permits him to make a rudimentary form of inference. If he is confronted with an object he can form and retain a concept specifying those of its attributes to which he has attended. If he now encounters some new object, he can compare it with the retained object. But as he can simultaneously retain two concepts at most, he must classify his concepts of the two objects as being either identically the same, or quite different: he cannot have a third concept of their similarities and differences. Therefore, if the new object has at least one of the attributes of the retained concept, he may expect to find that it has all the other attributes as well. Stern (1924) calls this kind of reasoning from one concrete instance to another ' transduction ' : it is exemplified by his daughter who, at the age of 3, knew a rope as something to support the family wash, so when she saw a rope strung along a balcony, she asked : " What's that hanging up there ? Something to pin clothes on ? "

Thinking at this stage is also characterised by the phenomenon of *irreversibility*. It is exemplified by an infant who is shown some brown beads and some white beads. He will correctly tell a questioner that the brown beads are wooden and that the white beads are wooden. But, if he is shown a box containing a larger number of brown than white beads, and he is asked whether it holds more brown beads or more wooden beads, he will insist that there are more brown beads. This would be predicted of an infant who cannot think of three concepts simultaneously : for, if he has already classified the white wooden beads as ' wooden ,' and the brown wooden beads as ' brown,' or even ' brown wooden,' he has not the capacity to restructure his concepts to form a third concept ' wooden of any colour.'

LEVEL 2. This level demands the capacity to process up to $2^2=4$ concepts simultaneously. It corresponds to the *stage of concrete operational intelligence.* This is defined by Piaget and Inhelder (1959) as the time, between 7 and 11 years of age, when the child learns how to use *concrete operations*, which are processes of restructuring concepts, that can be illustrated by grouping tangible objects.

Suppose that a child has formed concepts of two classes, A and A′. If he considers the attributes of A and A′ simultaneously, he thereby forms the concept of the sum of A and A′. This new third concept may be designated B. Now the child's immediate memory can accommodate the concept of a further class, B′, which can in turn be coupled with the concept of B. This process, called the concrete operation of *addition*, may be continued indefinitely. It is manifested when the child couples classes of objects to form a wider class. Addition is also manifested when the child puts objects into a series ordered by some specific relation. This is now possible because the child has the capacity, not only to retain concepts of each pair of objects involved,but also the further concept of the relations between them.

The notion that the child forms wider class concepts by combining concepts of subclasses two at a time is similar to that advanced by Piaget (1942 ; 1947, p. 43 ; 1949), who states that a child forms the concept of a class (e.g., animals) by combining the concept of a subclass (e.g., insects) with the concept of the complement of that subclass (non-insects). But the complement seems to be defined *with respect to the new class* (thus making ' non-insects ' mean ' *animals* that are not insects '). This would imply that the child must already have the concept of the new class in order to form that concept—a circularity which obviously cannot be intended.

Next suppose that the child has in his immediate memory the concepts of two classes C and D. Then the child can classify further concepts using, at the same time, the two criteria, ' having the attributes of C ' and ' having the attributes of D.' This is manifested in the concrete operation of *multiplication*, that is, in sorting objects into four product classes according to whether they have the attributes of C and D ; of C but not of D ; of D but not of C ; or of neither C nor D. Multiplication is also manifested when a child finds correspondences between two given independent series.

The delay between the attainment of a four-chunk memory span, between the ages of 4 and 5, and the start of the concrete operational stage, at about 7 or 8 years, is more apparent than real. In fact, Piaget and Inhelder (1959) give statistical tables which may be summed up by saying that no 4-year-olds can perform concrete operations, but that some can be carried out systematically by an increasing number of subjects from the age of 5. Evidence of the early performance of concrete operations is also given by Welch (1940) who, incidentally, shows the relation between the abstractness of concepts and their number. Welch distinguishes stages in the development of abstract thought by the hierarchic order of concepts a child can form. At 26 months, on average, his subjects attained first-order concepts: for example, potatoes and apples were both classed as foods. The second stage was reached at 4½ years, when his subjects attained second-order concepts, potatoes being classed as vegetables, apples as fruit, and both fruit and vegetables as foods.

Concrete operations are said to be *reversible*, because the child can behave as if he had mentally cancelled an operation, either by doing the opposite (a type of reverse operation called *inversion*), or by performing an operation that compensates for the first one (*reciprocity*). Thus, for the first time he can argue that if lemonade is poured from a tall, slim glass into a short, fat one, there will still be the same amount to drink, because the lemonade can be poured back again to look as it did at first (inversion), and because decreasing the lemonade's depth has been compensated for by giving it more breadth (reciprocity).

At this stage, however, the child cannot understand how inversion and reciprocity can be combined and so effect *correlation*. For example, the child may be faced with an apparatus consisting of two communicating vessels containing a liquid; if a piston is pressing on the liquid in one vessel, the level of the liquid in the other vessel will vary according to the weight of the piston and the density of the liquid. A child, at this stage, cannot educe the correlate needed to answer the problem: " The operation of increasing the weight of the piston is to decreasing the weight of the piston (inversion) as increasing the density of the liquid (reciprocity) is to . . . what? " (Inhelder and Piaget, 1955, p. 148 ff.). The hypothesis of an immediate memory limited to holding four concepts at a time accounts for the child's inability to educe such a correlate, for to understand the problem he would have to retain not only the concepts of an operation, its inversion and its reciprocity, but also concepts of the three relationships involved.

The hypothesis also explains why it is not until the next stage that a child can remember to keep all factors constant except one during an investigation (Inhelder and Piaget, 1955, p. 46 ff.).

LEVEL 3. This level demands the ability to process up to $2^3=8$ concepts simultaneously. This corresponds to the *stage of formal operational intelligence*, which is reached by 11 to 15 years, when the adolescent thinks in conformity with a fully practical logic.

If the adolescents can now retain up to eight concepts, he can, in principle, make any inference involving classes distinguished by up to three attributes (or sets of attributes), such as a syllogism, so the limitations of the previous level vanish for nearly all practical purposes.

Indeed, it will be an unusally complex problem that demands the retention of as many as eight concepts. The conclusion to the simplest form of syllogism (known as Barbara) can be found even with level 2 logic by a process akin to logical addition: the first premiss says that all of class A has the attributes of class B, so that the concept of their sum, A and B, can be formed and used to substitute for A or for B; the second premiss says that all of class B has the attributes of class C, so that class A has the attributes of class C, which is the required conclusion. However, if this kind of reasoning is applied to other premisses, it may produce invalid conclusions. In fact, most fallacies in reasoning can be accounted for on the hypothesis that the problem being considered calls for the construction of more concepts than can be kept distinct in the immediate memory.

A level 4 logic, which would enable subjects to solve problems involving concepts varying in four different ways, is, of course, possible. However, it is a rare person who is intelligent enough to solve such problems mentally, because they require the simultaneous retention of up to $2^4=16$ concepts.

ACKNOWLEDGEMENTS.—Thanks are due to Professor E. A. Peel and Dr. A. R. Jonckheere for sugggesting improvements to an earlier draft of this paper.

REFERENCES.

(Page numbers in the text refer to English translations where possible)

BERNSTEIN, B. A. (1935). On finite Boolean algebras. *Amer. J. Math.*, 57, 733-742.

BETH, E. W. (1950). A propos d'un " Traité de logique." *Methodos*, 2, 258-264.

GRIZE, J. B. (1960). In *Etudes d'épistémologie génétique, XI—Problèmes de la construction du nombre*. Ed. J. Piaget. Paris : Presses universitaires.

INHELDER, B., and PIAGET, J. (1955). *De la logique de l'enfant à la logique de l'adolescent.* Paris : Presses universitaires. Trs A. Parson and S Milgram (1958) *The Growth of Logical Thinking*. London : Routledge and Kegan Paul.

LANGER, S. K. (1953). *An Introduction to Symbolic Logic*, 2nd ed. New York : Dover.

LEWIS, C. I., and LANGFORD, C. H. (1932). *Symbolic Logic*, 2nd ed. New York : Dover

MILLER, G. A. (1956). The magical number seven, plus or minus two : some limits on our capacity for processing information. *Psychol. Rev.*, 63, 81-97.

MUNN, N. L. (1954). In *Manual of Child Psychology*. Ed. L. Carmichael New York : Wiley.

PIAGET, J. (1937). *La construction du réel chez l'enfant.* Neuchatel : Delachaux et Niestle. Trs. M. Cook (1955). *The Child's Construction of Reality.* London : Routledge and Kegan Paul.

PIAGET, J. (1942). *Classes, relations et nombres.* Paris : Vrin.

PIAGET, J. (1947). *La Psychologie de l'intelligence.* Paris : Colin. Trs. M. Pearcy and D. E. Berlyne (1950). *The Psychology of Intelligence.* London : Routledge and Kegan Paul.

PIAGET, J. (1949). *Traité de logique . Essai de logistique opératoire.* Paris : Colin.

PIAGET, J. (1952). *Essai sur les transformations des opérations logiques.* Pâris : Presses universitaires.

PIAGET, J. (1953). *Logic and Psychology.* Manchester : University Press.

PIAGET, J. (1957). In *Etudes d'épistémologie génétique, II—Logique et équilibre.* Ed. J. Piaget. Paris : Presses universitaires.

PIAGET, J., and INHELDER, B. (1941). *Le développement des quantités chez l'enfant.* Neuchatel : Delachaux et Niestle.

PIAGET, J., and INHELDER, B. (1959). *La genèse des structures logiques élémenataires.* Neuchatel : Delacahux et Niestle.

STERN, W. (1924). *Psychology of Early Childhood.* New York : Holt.

WECHSLER, D. (1949). *Wechsler Intelligence Scale for Children* (Manual). New York : The Psychological Corporation.

WECHSLER, D. (1955). *Wechsler Adult Intelligence Scale* (Manual). New York : The Psychological Corporation.

WELCH, L. A. (1940). A preliminary investigation of some aspects of the hierarchical development of concepts *J. Gen. Psychol.*, 22, 359-78.

(*Manuscript received, 9th September,* 1962.)

PREDICTING STUDENT PERFORMANCE IN A UNIVERSITY DEPARTMENT OF EDUCATION

By F. W. WARBURTON, H. J. BUTCHER AND G. M. FORREST
(Department of Education, University of Manchester)

SUMMARY. One hundred teachers in training in the Department of Education, University of Manchester, were tested with measures of abilities, personality, interests, values and general culture, and provided a considerable range of biographical information. The predictive power of these various measures was examined in terms of three main criteria—final teaching mark, final theory mark, and award of the Graduate Certificate in Education. Degree class was the best single predictor of theory mark and of final award. A personality questionnaire (Cattell's 16 P.F.) was the best single predictor of teaching mark. The successful student tended, apart from academic qualifications, to have a high level of general culture, high conscientiousness (16 P.F., factor G), to have tender-minded attitudes to education, and to participate fairly widely in social activities.

I.—INTRODUCTION.

THE literature dealing with the prediction of success in teaching and success in teachers' training courses is very large and very inconclusive. Some idea of its extent may be obtained from the annotated biography of Domas and Tiedeman (1950) which contains over 1,000 references to papers on ' teacher competence ' up to 1950. Its inconclusiveness is also well known. As Evans (1959) says: " The most obvious deduction from the available evidence is that there is no one pattern of successful teacher. Teaching is a complex process calling for many different abilities, no one of which is by itself sufficient to ensure success. At the same time, each of these abilities makes a small but necessary contribution to success " (p. 33).

Among the variables which have been previously studied as possibly related to teaching success are: biographical data such as age, marital status (e.g., Ryans, 1951), intelligence and abilities (in a large number of studies), personality factors measured by questionnaire or by objective tests (Lamke, 1951; Schwartz, 1950), interests (Strong, 1943), attitudes to teaching (Cook, Leeds and Callis, 1952) and many others. In addition to the study of theoretical predictors of this type, much research has been devoted to the predictive power of interviews and kindred methods of selection (e.g., Burroughs, 1958).

The results have, in general, been rather disappointing. Several causes have clearly contributed to the prevailing uncertainty in this field. Firstly, as stated by Evans in the quotation above, the criterion is a complex one. Secondly, in some studies the criterion has been teaching mark at student level, in others rated success as a teacher after varying periods of service. Little is known about the relation between these two criteria, as is pointed out by Collins (1959), since the few correlations available range from +·84 (Tudhope 1942) to +·16 (Martin, 1944). Thirdly, the numerous variables which have been tested as predictors have generally been tried out on quite different samples, varying in age, nationality, sex, degree of selection and other respects, with the result that it is hard to compare the results obtained and the relative values of the different types of predictor.

These last two factors are our justification for the present paper, for it describes the administration of tests and the collection of data of all the types described above with one sample of students in a typical university department of education. Although the present research uses the teaching practice marks of graduate students as its principal criterion, it is also planned to follow up the sample after five years' teaching experience, and thus, firstly, to provide further evidence about the relation between practice marks and later success, but also, more importantly, perhaps, to determine which of the predictors of student teaching success hold up, and which deteriorate, when considered against the later and more fundamental criterion of success in teaching as a career.

II.—DETAILS OF THE RESEARCH.

(a) *Objectives.*

The programme of research thus has three main aims : (i) to obtain as complete a picture as possible of a sample of graduates entering the teaching profession in terms of biographical and academic background, abilities, interests, values and personality factors ; (ii) to determine the relative efficiency of these variables in predicting success in the course of professional training ; (iii) to follow up the subjects after five years' teaching experience and re-determine the relative predictive values in terms of success in teaching as a career. The present paper is mainly concerned with the second of these aims.

(b) *The subjects.*

Students taking the Graduate Certificate in Education course in the Department of Education, University of Manchester, in the session 1957-58, numbered 133. Volunteers were called for to undergo an extensive programme of psychological testing, and 118 attended the first period of testing, but inevitably by the time data on approximately 100 variables had been assembled, a number of gaps had appeared. The number of students on whom information on every variable was available had shrunk to rather less than 100. However, to facilitate analysis by the electronic computer, and to avoid the complication of slightly different numbers in the sample in different variables, the gaps were filled by assigning to a student who had missed a test, or about whom information was missing, the mean score of the whole group on that particular variable. This expedient was necessary in 2·3 per cent. of the entries in a 100 × 100 score matrix.

(c) *The variables.*

The variables included in the research comprise fifty-seven test scores, twenty-five biographical particulars and eighteen criteria of success on the training course, as follows :

TABLE 1

(1) TEST SCORES.

1—Non-verbal intelligence	Culture Free Scale 3A (Cattell).
2—Verbal Intelligence.	National Institute of Industrial Psychology. Group Test 33
3—Spatial ability.	National Institute of Industrial Psychology. Group Test 80A
4—Memory.	Primary Mental Abilities Thurstone.
5—Verbal Fluency.	Primary Mental Abilities. Thurstone.
6—Mechanical Comprehension.	Mechanical Comprehension. Bennett.

7—Vocabulary.	Primary Mental Abilities. Thurstone (70 items), together with 30 difficult words from various other sources.
8—General Culture.	100 item test, constructed by H. J. Butcher and F. W. Warburton, Department of Education, University of Manchester. The subject is required to give the name of the authors of a list of books, covering the humanities, sciences and social sciences and works of art and music.
9—Cyclothymia *vs.* Schizothymia.	16 P.F. Factor A (Cattell)
10—Intelligence.	,, ,, Factor B ,,
11—Ego Strength.	,, ,, Factor C ,,
12—Dominance.	,, ,, Factor E ,,
13—Surgency.	,, ,, Factor F ,,
14—Conscientiousness.	,, ,, Factor G ,,
15—Extraversion.	,, ,, Factor H ,,
16—Sensitivity	,, ,, Factor I ,,
17—Paranoid Tendency	,, ,, Factor L ,,
18—Unconventionality.	,, ,, Factor M ,,
19—Sophistication.	,, ,, Factor N ,,
20—Guilt Proneness (Anxiety)	,, ,, Factor O ,,
21—Radicalism	,, ,, Factor Q_1 ,,
22—Self-sufficiency.	,, ,, Factor Q_2 ,,
23—Self-control.	,, ,, Factor Q_3 ,,
24—Id Pressure (Tenseness)	,, ,, Factor Q_4 ,,
25—Outdoor Interests.	Kuder Preference Record–Vocational Form CH.
26—Mechanical Interests.	,, ,, ,, ,,
27—Computational Interests.	,, ,, ,, ,,
28—Scientific Interests.	,, ,, ,, ,,
29—Managerial Interests.	,, ,, ,, ,,
30—Artistic Interests.	,, ,, ,, ,,
31—Literary Interests.	,, ,, ,, ,,
32—Musical Interests.	,, ,, ,, ,,
33—Social Interests	,, ,, ,, ,,
34—Clerical Interests.	,, ,, ,, ,,
35—Theoretical Values.	Study of Values (Allport, Vernon and Lindzey).
36—Economic Values.	,, ,, ,, ,,
37—Aesthetic Values.	,, ,, ,, ,,
38—Social Values	,, ,, ,, ,,
39—Political Values.	,, ,, ,, ,,
40—Religious Values	,, ,, ,, ,,
41—Radical attitudes.	Inventory of Social Attitudes (Eysenck).
42—Tenderminded attitudes	,, ,, ,, ,,
43—Radical attitudes to education.	Survey of Opinions about Education (Oliver), Department of Education, University of Manchester.
44—Tenderminded attitudes to education.	,, ,, ,, ,,
45—Naturalist attitudes to education.	,, ,, ,, ,,
48—Accuracy in non-verbal intelligence test	Proportion answers correct in Test 1.
49—Accuracy in verbal intelligence test.	Proportion answers correct in Test 2.
50—Accuracy in spatial ability test.	Proportion answers correct in Test 3.
51—Accuracy in mechanical ability test.	Proportion answers correct in Test 6.
52—Accuracy in vocabulary test.	Proportion answers correct in Test 7.
53—Accuracy in general culture test.	Proportion answers correct in Test 8.
54—Emphasis in personality test.	No. 'extreme' answers in 16 P.F. Test.
55—Emphasis in values test.	No. extreme answers in study of Values test.
56—Emphasis in social attitudes test.	No. extreme answers in Inventory of Social Attitudes.
57—Emphasis in radical attitudes to education.	No. extreme answers (radical) Survey of Opinions about Education.

58—Emphasis in tender-minded attitudes to education.	No. of extreme answers (tender-minded) Survey of Opinions about Education.
59—Emphasis in naturalist attitudes to education.	No. extreme answers (naturalistic). Survey of Opinions about Education.

(2) BIOGRAPHICAL PARTICULARS.

60—No. of G C.E. 'O' or 'A' subjects.	No. of subjects passed either at 'O 'or 'A' level in General Certificate of Education, i.e., subjects in which at least 'O' level is reached.
61—No. of G.C.E. 'A' subjects.	No. of subjects passed at 'A' level in General Certificate of Education.
62—Degree class.	First class honours=5. Second class honours (upper division)=4. Second class honours (lower division)=3. Third class honours=2 Pass degree=1. No degree=0.
63—Grant.	State scholar=2. Otherwise grant aided=1. No grant=0.
64—Previous teaching experience.	Questionnaire. Any form of teaching, e.g., supply, Sunday School, armed forces=1. No teaching=0.
65—Degree course tutor's reference.	A–E scale.
66—Interview grade.	Mean grade of 2–3 individual interviewers working separately. A–E scale.
67—Social status of school.	Independent School=2. Direct Grant=1. Grammar School=0.
68—Age of university.	Classified as follows: Oxford=2. Cambridge=2. London=1. Manchester=1. Liverpool=1. Leeds=1. Sheffield=1. Nottingham=1. Durham=1. Southampton=0. Exeter=0. North Staffordshire=0. Aberystwyth=0.
69—Arts Faculty.	Arts=2. Social Science=1. Science=0.
70—Age.	
71—Male Sex.	
72—No. of University offices.	Questionnaire. No. given.
73—No. of University societies.	Questionnaire. No. given.
74—No. of interests.	Questionnaire. No. given.
75—No. of games.	Questionnaire. No. given.
76—University residence.	University Hall of Residence=2. Lodgings=1. Home=0.
77—Father's education.	University=5. Training College=4. Public School=3. Grammar School, High School=2. Technical School, Commercial School, Trade School=1. Elementary School=0.
78—Mother's education.	As 'Father's education.'

79—Parent's occupational level.	Grades I–V. Classification of Occupations (H.M.S.O.).
80—Large family size.	No. of children.
81—Late position in family.	1st, 2nd, etc
46—Religious affiliation.	Questionnaire Fundamentalist Sects=5. Nonconformists=4. Roman Catholics=3. Church of England=1. No religion=0.
47—Right political affiliation.	Questionnaire. Conservative=4. Liberal=3 Labour=1. Communist=0.
99—Entry grade.	A–E scale A mark given before entrance to the Department of Education. It is based on the applicant's complete record, including examination, tutor's reference, interview grades and life record.

(3) CRITERIA OF SUCCESS.

91—Final Teaching Mark.	A–E scale. Graduate Certificate in Education examination.
97—Final Theory Mark	,, ,, ,, ,,
98—Final Certificate Award.	,, ,, ,, ,,
82—Tutor's Teaching grade (Michaelmas Term)'	A–E scale by tutor supervising teaching practice
83—Tutor's teaching grade (Lent Term).	A–E scale by tutor supervising teaching practice.
84—Schools' teaching grade (Michaelmas term)	A–E scale by headmaster of school in which student's school practice was carried out.
85—School's teaching grade (Lent term).	,, ,, ,, ,,
86—Tutor's teaching check list (Michaelmas term).	Check list of 43 items concerned with student's disciplinary control, lesson construction, attitude to pupils, etc.
87—Tutor's teaching check list (Lent term)	,, ,, ,, ,,
88—Extra Tutor's teaching check list.	,, ,, ,, ,,
89—School's teaching check list (January).	,, ,, ,, ,,
90—School's teaching check list (April).	,, ,, ,, ,,
92—Psychology of education examination	A–E scale Graduate Certificate in Education examination.
93—Philosophy of education examination.	,, ,, ,, ,,
94—History of education examination.	,, ,, ,, ,,
95—Health education examination	,, ,, ,, ,,
96—Essay.	,, ,, ,, ,,
100—Curriculum and Method examination	,, ,, ,, ,,

Scores on the 100 variables were transformed into normalized T scores. The resulting 4,950 product moment correlations were calculated, on the Manchester University electronic computer. Several factor analyses were also carried out, which it is hoped to report in a later paper.

(*d*) *Particulars of the main criteria.*

The prediction of teaching success to be reported in the next sections will be in terms of three criteria, (*a*) final teaching mark, (*b*) final theory mark, (*c*) award or non-award of the certificate. All three refer to the Graduate Certificate in Education of the University of Manchester. In general, the process of arriving at these marks and at the award is believed to be fairly typical of the procedure in other university departments of education.

Final teaching mark is based primarily on a tutor's observations and assessment of a student during about twelve weeks of school practice, although reports from the schools and from an external examiner will, in many cases, be bound to affect it. Final theory mark is based on the average of five examination papers and one long essay. Award of the Graduate Certificate represents the decision of a group of examiners on receiving reports of final teaching mark and final theory mark, together with the testimony of examiners who know the student.

III.—THE PREDICTION OF SUCCESS AT TEACHING.

(*a*) *Prediction from individual measures.*

All significant correlations obtained between any of the tests or biographical variables and the three main criteria are shown in Table 2 below :

TABLE 2

SIGNIFICANT CORRELATIONS BETWEEN THE PREDICTORS AND THE MAIN CRITERIA OF SUCCESS

Predictors	Criteria		
	Final Teaching Mark	Final Theory Mark	Certificate award
TESTS .			
7—Vocabulary	—	·236	—
8—General Culture	—	·399*	·308*
11—Ego strength (16 P.F)	—	·194	—
14—Conscientiousness (16 P F.)	·236	·236	198
16—Sensitivity (16 P F)	·220	—	·215
18—Unconventionality (16 P F)	—	·199	—
23—Self-Control (16 P F)	·223	—	—
36—Economic Values (Vernon)	—	—·213	— 257*
44—Tendermindedness in education	—	·199	·210
49—Accuracy (verbal intelligence)	—	·203	—
52—Accuracy (vocabulary)	—	·258*	·231
54—Extreme responses to personality questionnaire	—	—·275*	—·204
57—Emphasis on scale of radicalism in education	—	·262*	·246
BIOGRAPHICAL PARTICULARS			
60—G C E 'O' or 'A'	—	·244	·305*
62—Degree class	·238	·569*	·514*
63—Grant	—	·257*	·231
65—Degree Course, Tutor's reference	·249	·366*	·415*
66—Interview grade	·265*	·289*	·254*
72—No. University offices held	—	·242	—
74—No of interests	—	·261*	·263*
99—Entry grade	·276*	358*	·322*

All the correlations in the above table are significant at the ·05 level, except those marked with an asterisk, which are at the ·01 level.

(i) *Prediction from Psychological tests.*

Table 2 shows that of the 171 correlations between test scores and the three main criteria, only twenty-two (13 per cent.) are significant. This may seem a small number, but it must be remembered that we would not necessarily expect a large proportion of variables to be related to success in teaching, or in any other particular profession. As regards the type of test, significant correlations were obtained for 17 per cent. of the cognitive test scores, 18 per cent. of the personality measures, 0 per cent. of the interests, 10 per cent. of the values and 15 per cent. of the attitudes assessments (22 per cent. for attitudes to education). There is an important reason, however, why the size of the correlation coefficients should not be taken quite at its face value. This is the familiar problem of correction for range. Although it is impossible to state the extent to which our sample of graduate students forms a selected group compared with the total group of successful and unsuccessful applicants for entry to the course, since the degree of selection and consequent narrowing of range will be different for every variable, there is little doubt that if the tests and predictors had been administered to all candidates, the majority of the correlation coefficients would have been markedly higher, and many of those that are at present only suggestive would become statistically significant. A striking example of the effects of correction for selection in a rather similar situation is provided by Vernon's (1950) study of Civil Service Selection procedure, in which all the separate, uncorrected validities are extremely low, but, when corrected for selection and combined, yield a considerable degree of prediction. On the whole, the results indicate the relative success of two single long tests, Cattell's 16 P.F. personality test, and Oliver's "Survey of Opinions about Education," and the relative failure of the cognitive tests and the measures of interests. For success on the theoretical side of the course, culture and vocabulary are important, suggesting that Arts students, with their literary and verbal background, have an advantage in this respect. Ego-strength, conscientiousness and lack of extreme personality characteristics are also important, perhaps in the preparation of examinations. There is also some indication that tender-mindedness, radical attitudes and originality (unconventionality) are a factor in examination success, i.e., in pleasing the examiners.

The 'General Culture Test' emerges as a better predictor of theory mark than any other single test. This unpublished test (Butcher and Warburton) is *prima facie*, a rather crude measure of extent of reading. It consists of 100 items, the large majority of these being the titles of books, with the respondent being required to name the author. In the remainder of the items, the same principle is applied to well-known paintings and musical compositions.

Turning to the variables that do *not* significantly predict theory mark, it is noticeable that all six of the ability tests (variables 1-6) fall into this class, presumably because of the high general intellectual level, and hence the narrow range, of the sample. It is worth noting that whereas one at least of each type of test correlates with the Final Theory Mark (and hence with the Certificate Award), only personality scores are related significantly to practical teaching. The finding that practical teaching is harder to predict than theoretical success is typical of research in this area. The personality qualities related to Practical Teaching are self-control (Q3), conscientiousness (G) and sensitivity (I). This finding is in agreement with an American research by Erickson (Barr, *et al.*, 1961), who administered the 16 P.F. to a group of sixty-four teachers in Wisconsin High Schools. Erickson found factors Q_3, G and M– to be related to teaching success, with obtained correlations which were all of the order ·2 to ·3. The first two qualities are aspects of the major second order personality factor of stability, while sensitivity is related to tendermindedness, an attitude which is shown throughout this study to be positively correlated with success on this training course. This is not a surprising result in view of contemporary trends in education.

Positive correlations between factors G and Q_3 and teaching success are reported in the 16 P.F. handbook, but the correlation with factor I (Emotional Sensitivity) is a new finding which may, however, be related to cultural differences between the U.S.A. and Britain. Certainly, both subjective observation and empirical research (e.g., Cattell and Warburton, 1961) suggest that this quality of sensitivity or tendermindedness is both more highly valued by and more typical of the European than of the North American culture-pattern. One value of the 16 P.F. test in the present context appears to be that it alone, among the wide range of tests used, predicts teaching practice better than teaching theory.

Besides the significant correlations, it is interesting to note those that, perhaps against expectation, are of negligible size. In particular, verbal and non-verbal intelligence are irrelevant to teaching success in this sample, except in so far as non-verbal intelligence shows a *negative* (but non-significant) correlation of —·125. Similarly, the extensive series of scores provided by the Kuder Preference Record and the Allport-Vernon measures of values fail to yield any significant correlation, although one or two are suggestive of some relationship.

Non-significant but suggestive correlations with teaching mark are shown below:

Age	+·173
No. of G.C.E. passes	+·172
No. of interests (as stated on application form for entry to course)	+·166
Family size	+·165
Accuracy on vocabulary test	+·165
Verbal fluency	+·155
Naturalism in education	—·151
Paranoia or suspiciousness (16 P.F.)	—·142
Social interests (Kuder)	+·138
Economic values (Allport-Vernon)	—·136

In general, these correlations, although small, make quite good psychological sense and require little explanatory comment. The concept of 'naturalism' in education (Butcher, 1959; Oliver and Butcher, 1962) refers approximately to a 'child-centred,' 'progressive' grouping of attitudes. A teacher scoring high on the scale of naturalism would be expected, for instance, to approve and follow many of the precepts of John Dewey.

(ii) *Prediction from biographical data.*

Twenty-five per cent. of the correlations between biographical particulars and the three main criteria are significant as compared with 13 per cent. for tests. As with the psychological tests, examination success is easier to predict than practical teaching. Academic capacity is important, but there is also a relationship with breadth of interests (number of interests and university offices held). Two points of interest may be mentioned about these results in respect of theory marks. Firstly, the correlation between degree class and theory mark is remarkably high, particularly when one considers the necessarily coarse grading in which degree class has to be expressed. From the point of view of predicting theory mark (and, indeed, award of the certificate), degree class is much more efficient than entry grade, although all available evidence is taken into account in arriving at the latter.

Secondly, this is also true of tutor's reference. Both theory mark and award of the certificate are better predicted by the report on a student's performance as an undergraduate than they are by entry grade, although the latter includes knowledge of the former. Degree class is already known to the interviewer, and entry grade is based on all the information available to the interviewer, on the results of the interview, and on any further information that may have subsequently become available, such as additional references. Practical teaching ability is related to four variables but these essentially comprise academic level only.

Putting the test scores and biographical data together, the good classroom teacher appears to be essentially stable, sensitive and of good academic capacity. Students successful in examination have the same characteristics, together with good verbal capacity and tenderminded and radical attitudes to educational problems.

(*b*) *Multiple Correlations.*

Turning to the multiple correlations (Table 3), we see that the personality variables are the best set of predictors for practical teaching, the theory examination and the final certificate award. These variables are all derived from a single test, Cattell's 16 P.F., which can be administered in 40 minutes. It is, of course, commonplace knowledge that personality is of a very considerable importance, and possibly of paramount importance, in determining academic capacity. The present results suggest that it is no longer necessary to neglect personality test scores in predicting student performance. Interest tests scores are also more efficient predictors than the eight ability tests. The test of general attitudes was understandably less effective than the test of educational attitudes. There is a tendency, but it is not very marked, for the multiple correlations to be higher for the theory examination than for practical teaching.

An interesting finding was that, on the whole, twenty-one biographical particulars afforded better prediction (·871) than forty-two psychological variables (·639). It should, of course, be noted that in practice beta weights on different samples tend to fluctuate wildly and multiple correlations fall considerably when the same weights are applied to later samples.

TABLE 3

MULTIPLE CORRELATIONS.

	Practical Teaching	Theory Examination	Final Certificate Grade
8 ability tests	·270	·466	·335
15 personality variables	·438	·538	·448
10 interest scores	·330	·293	·268
6 values	·167	·265	·298
2 attitudes	·084	·170	·134
3 educational attitudes	·207	·241	·257
42 psychological variables	·653	·696	·639
21 biographical particulars	·501	·741	·871

IV.—IMPLICATIONS FOR SELECTION PROCEDURES.

Although considerations of space prevent a full discussion of the several factor analyses carried out, it will be convenient simply to list the seventeen main factors which emerged in the most satisfactory of the analyses and to discuss the most promising predictive measures under these headings. One factor loaded almost exclusively on criterion variables and was interpreted as teaching ability. The remaining sixteen were interpreted as follows :

(1) Academic success.*
(2) Verbal ability.
(3) Non-verbal ability.
(4) Stability of personality.*
(5) Extraversion.
(6) Radicalism.
(7) Tender-mindedness.*
(8) Emphasis in expressing opinions.
(9) Interest in arts subjects *versus* science subjects.
(10) Aesthetic interests.
(11) Participation in social activities.*
(12) Socio-economic status.
(13) Family size.
(14) Assessment on entry to Department.*
(15) Verbal fluency.
(16) General culture.*

Six of these factors were loaded by variables which appeared to have a promising relation with the main criteria, and these have been marked by asterisks in the above list. They are also shown, with the more promising test and biographical variables, in Table 4 below, which includes the combined predictive effects of the groups of variables. The correlations would probably be appreciably higher if the technique of multiple correlation had been used in place of the simple pooling with unit weights.

TABLE 4

TESTS TO BE USED IN SUGGESTED SELECTION PROCEDURE.

	Theory Marks	Practical Teaching	Certificate Award
(1) ENTRY ASSESSMENT.			
Final entry grade	+358	+276	+322
Tutors' reference (degree course)	+366	+249	+415
Total (unit weights)	+408	+296	+416
(2) ACADEMIC SUCCESS.			
University degree class	+569	+238	+514
Total (unit weights)	+569	+238	+514
(3) GENERAL CULTURE.			
General Culture (Butcher and Warburton)	+399	+005	+308
Vocabulary (Thurstone)	+236	+054	+156
Total (Unit weights)	+265	+128	+299
(4) STABILITY.			
Conscientiousness (16 P.F.G.)	+131	+236	+198
Self-control (16 P.F.—Q3).	−101	+223	+063
Total (Unit weights)	+018	+281	+160
(5) TENDER-MINDEDNESS.			
Sensitivity (16 P.F.—I).	+156	+220	+215
Tendermindedness in Education (Oliver)	+048	+199	+210
Total (Unit weights)	+231	+179	+277
(6) SOCIAL ACTIVITIES.			
No. of expressed interests	+261	+166	+263
No. of University offices held	+242	−029	+108
Total (unit weights)	+313	+086	+231
ALL TESTS (unit weights)	+707	+520	+708

In considering these findings, it is also interesting to note the factors to which success in the teaching course is largely unrelated. These are verbal and non-verbal ability, extraversion, radicalism, emphasis in expressing opinions, arts *versus* science interests, aesthetic interests, socio-economic status, family size and verbal fluency.

It would, therefore, appear best in selecting potential teachers to concentrate on attainment and general culture rather than ability, on stability rather than on extraversion or introversion, on academic rather than aesthetic interests, and on participation in social activities rather than on social or domestic background.

V.—REFERENCES.

BARR, A. S., *et al.* (1961). Wisconsin studies of the measurement and prediction of teacher effectiveness. *J. Exper. Educ.*, 30, 1-156.

BURROUGHS, G. E. R. (1958). A study of the interview in the selection of students for teacher training. *Brit. J. Educ. Psychol.*, 28, 1, 37-46.

BUTCHER, H. J. (1959). The attitudes of teachers and student teachers to education. *Unpublished Ph.D. thesis*, Univ. of Manchester Library.

CATTELL, R B. (1957). *Personality and Motivation Structure and Measurement.* New York. World Books

CATTELL, R. B., and WARBURTON, F. W. (1961). A cross-cultural comparison of patterns of extraversion and anxiety. *Brit. J. Psychol.*, 52, 1, 3-15.

COLLINS, M (1959). A follow-up study of some former graduate student teachers. *Brit J. Educ Psychol.*, 29, 3, 187-197.

COOK, W. W., LEEDS, C. H., and CALLIS, R. (1952) Manual *of the Minnesota Teachers' Attitude Inventory*. New York : Psychological Corporation.

DOMAS, S J., and TIEDEMAN, D. V. (1950). Teacher competence : an annotated bibliography. *J. Exp. Educ.* 19, 2, 101-218.

EVANS, K. M. (1959). Research on teaching ability. *Educ Research* 1, 22-36.

LAMKE, T. A. (1951). Personality and teaching success. *J. Exper. Educ.*, 14, 75-100

MARTIN, L. O. (1944). *The Prediction of Success for Students in Teacher Education* New York : Teachers' College, Columbia University.

OLIVER, R. A. C., and BUTCHER, H. J. (1962). Teachers' attitudes to education. I.—The structure of educational attitudes. *Brit. J. Soc Clin. Psychol.*, 1, 56-69.

RYANS, D G. (1951). A study of the extent of association of certain professional and personal data with judged effectiveness of teacher behaviour *J. Exp Educ.*, 20, 1, 67-77.

SCHWARTZ, A. N. (1950). A study of the discriminating efficiency of certain tests of the primary source personality traits of teachers *J Exp Educ.*, 19, 1, 63-94.

STRONG, E. K. (1943). *The Vocational Interests of Men and Women.* Stamford · Stamford University Press.

TUDHOPE, W. B. (1942 and 1943). A study of the training college final teaching mark as a criterion of future success in the teaching profession. *Brit J. Educ. Psychol.*, 12, 167-171 and 13, 16-23.

VERNON, P. E. (1950). The validation of civil service selection board procedures. *Occup. Psychol.*, 24, 75-95.

(*Manuscript received, September, 1962.*)

CRITICAL NOTE

SPECIFIC DYSLEXIA

BY D. RUSSELL DAVIS* AND ASHER CASHDAN†
(*Medical Psychology Laboratory, University of Cambridge*)

In areas of psychopathology where aetiologies are but imperfectly understood, and where methods of treatment are still a matter of argument, it is not surprising that controversy should rage over appropriate groupings of symptoms and the terminology most suitable for them. One such case is that of specific developmental dyslexia (formerly word-blindness), which under its new name has recently been receiving fresh attention (see Zangwill, 1960 ; Vernon, 1962 ; Miles, 1962) In our opinion, the controversy over the existence of this condition is often based on an imperfect understanding of the cardinal principles of scientific investigation and classification.

A disquietingly large number of children suffer from serious backwardness in reading Some catch up quickly, others never learn to read well. Some respond to specialist remedial teaching , others show a temporary improvement, but relapse as soon as the special teaching is withdrawn (see Collins, 1961) ; others show no positive response at all Some of the sufferers are girls, but more of them are boys. Some are of normal or superior intelligence, far more of them are of low intelligence. Many of them have other educational disabilities ; many also have emotional and family problems. A few of them have unusual patterns of handedness and eye dominance ; a few also are claimed by some investigators to have special perceptual difficulties, or difficulty in associating symbols ; such perceptual difficulties may be limited to the activity of reading, or may be more general.

The above list is not exhaustive, but implicit in it is the problem that concerns us. This is, to what extent are any, or all, of these factors associated with each other. If they are, we need to know which of the associations are causal ones, and in which direction the causal relationship operates. For instance, does backwardness in reading produce emotional disturbance, or do emotional disturbances lead to a refusal to learn to read ? Or are both true ?

If we studied and treated each case of reading disability separately, we might achieve fair success in curing the disability (or removing the impediment to learning). We would not, however, be able to draw any conclusions about the cause of the disability or the mode of operation of the ' cure ,' unless we were able to plan and carry out the treatment experimentally—and a single case is rarely sufficient for a satisfactory experiment. Again, if we gave identical remedial treatment to two groups of children, each having distinct features—such as right-handed and left-handed children, or children of low and high tested intelligence, or simply boys and girls—suppose then that one group improved significantly more than the other, we might conclude that left-handed children, say, have more persistent difficulties than right-handed ones, or that girls recover more easily than boys Even in such a situation, alternative hypotheses would be tenable, but these alternatives could reasonably be investigated However, the fact that one group had done better than another would have justified the classification (provided always that the result was a stable one). In other words, it would be reasonable to classify retarded boys separately from retarded girls if the two groups responded differently to treatment. But, if the response of the two groups was not clearly different from each other, no case would so far have been made out for classifying them as separate groups of retarded readers.

* Now at the Department of Mental Health, Bristol University

† Now at the Department of Education, Manchester University

In fact, if any category is to be established within the group of backward readers, it must, to serve any useful purpose, be demonstrably different in a significant respect from other (or some particular group of) backward readers. The purpose of such a differentiation would be to make easier the task of preventing, curing, or forecasting the course of the condition. The study of symptoms, and their classification, is only useful in so far as it throws light on psychological mechanisms, or is somehow relevant to these three tasks

If a new category is proposed, we are suggesting then that it must satisfy one of the following criteria. It must be different in respect of aetiology, prognosis, or treatment indicated. If the new category differs only in respect of symptoms, then the differences are either trivial or point to a need for further study of the relevant aetiology, prognosis or treatment.

Let us now, in this light, examine the proposed category of specific dyslexia, and taking each criterion in turn, ask whether the members of this proposed category can be differentiated from other backward readers.

Little is known of the aetiology of any form of reading backwardness. There is almost certainly some relationship with low intelligence, and with social class, but even this is not well documented. Hereditary factors have been suggested by some investigators, but there is no clear-cut evidence. Hallgren (1950) found hereditary factors in almost all cases ; this finding is so divergent from those of other workers as to make us hesitant to accept any of his conclusions. Zangwill (1960) found a significant number of affected relations in the families of dyslexics with handedness problems. But his work is also inconclusive. Such studies raise, incidentally, the question of the value of history-taking to establish familial features in pathological groups. Quite simply, if the mother of a child with some pathology is invited to produce relatives with similar histories, she is highly inclined to do this (and at the same time to alleviate her own guilt feelings). The mother of a normal child has no such inclination. An investigation based on history-taking may be of value, however, if it is designed deliberately to show whether the pattern of incidence of affected relatives conforms to the one predicted from a particular hypothesis. Hallgren's hypothesis, for instance, was that the pathogenic gene is dominant, but his method of proving that the observed pattern of incidence does not depart significantly from the predicted pattern was hardly satisfactory.

So much for aetiology We can do no better on prognosis As Vernon (1962) has pointed out, we certainly need to know why, in groups of children given remedial teaching, some improve while others do not, but in fact we do not yet know why this happens. Similarly, too, with methods of treatment. Miles (1962) says that " there is broad agreement as to the general type of remedial teaching that is normally called for," and he advocates, amongst other things, emphasis on phonic rather than on look-and-say methods, and the use of all relevant sense-modalities in the learning process. We would agree with his prescription, but still await evidence that it is indicated for any *specific* group of poor readers.

We cannot, therefore, agree that there is any good evidence for Vernon (1962)'s assertion that " there does exist a number of cases who have great difficulty in learning to read, which difficulty seems not to be attributable to environmental factors or to emotional maladjustment." There are many reasons why psychologists and others might wish to spread such a belief. Some of these are related to the clinical management of cases, and others to the desire to support *a priori* theories about neurology and genetics. But there is not much evidence to support the belief.

There are two legitimate ways in which advances may be made in this field. One is by clinical investigation and by testing theories and hypotheses in the clinical situation, or even by developing new insights there. But this has to be followed by the second method, which is to examine the relevance of separable factors (whether aetiological, symptomatic or prognostic), *one at a time*, in specially selected groups of children who are backward in reading.

To sum up, it is legitimate to adopt for research purposes the hypothesis that there is a special class of backwardness in reading, which we may call specific developmental dyslexia, and which has one or more characteristics, and then to carry out experiments to test the hypothesis itself, or, more likely, a deduction from it. Whether such a hypothesis is a promising one at the present time is a matter of opinion. But it would only be legitimate to assert or imply that the existence of this class had been established if evidence could be provided of a quality which deserved general acceptance. We have indicated in this note the sort of evidence which would be required, and we have given our opinion that it has not yet been provided.

REFERENCES

COLLINS, J E. (1961) The effects of remedial education *Educational Monographs, University of Birmingham, Institute of Education, No* IV.

HALLGREN ,B (1950). Specific dyslexia *Acta Psychiat. Neurol Suppl* , 65.

MILES, T R. (1962) Specific dyslexia *Bull. Brit. Psychol., Soc.*, 48, 51-2.

VERNON, M D (1962). Specific dyslexia *Brit. J. Educ. Psychol* , 32, 143-150

ZANGWILL, O L (1960) *Cerebral Dominance and its Relation to Psychological Function.* Edinburgh Oliver and Boyd.

SPECIFIC DYSLEXIA

A REPLY TO DR. RUSSELL DAVIS AND MR. ASHER CASHDAN

By M. D. VERNON

It was made clear in my paper, " Specific Dyslexia," that the classification of certain cases of reading backwardness as ' specific dyslexics ' was presented as a hypothesis based upon the association together of certain symptoms and types of behaviour with reading backwardness with above chance frequency, and in the absence of certain other characteristics such as low intelligence and emotional maladjustment (except in so far as the latter developed after rather than before the onset of reading difficulties). There seems to be reliable evidence to support this hypothesis, notably that obtained by Dr. Ingram (to whom Dr Davis and Mr. Cashdan do not refer). This hypothesis does no more than suggest tentatively what might be the causal factors, which are obviously very obscure. But it is, I suggest, scientifically interesting, and valuable in so far as it may stimulate further research (which may, of course, disprove it). Whether the hypothesis is useful in " the task of preventing, curing or forecasting the course of the condition " is a medical question on which I am not qualified to adjudicate ; though I would at least suggest that it might be desirable to consider whether backward readers who might be ' specific dyslexics ' should be treated in a somewhat different manner from those whose backwardness appears to be due primarily to environmental factors. Incidentally, I think that the term ' specific dyslexic ' may be preferable to that of ' specific dyslexia ' in so far as the latter suggests the existence of a single well-defined clinical entity, which seems improbable.

I entirely agree that the clinical diagnosis of cases of reading backwardness is often woefully inefficient, and I have said so repeatedly. I have also urged that much more thorough and systematic experimental examination is necessary of the psychological processes characterizing backward readers, as well as of their clinical histories. I agree also that all too little is known about prognosis and treatment. But I cannot agree that " the second legitimate way of advance " must be to examine the relevance of separable factors *one at a time*. This method has been utilized again and again in the past with little useful result. What we need to observe now is the nature and frequency of the association between certain aetiological, symptomatic and prognostic factors in backward readers. When these have been established, it may be possible to offer further hypotheses as to causality.

In support of what I have said here, I should like to quote from a paper in a recently published book, *Reading Disability* : *Progress and Research Needs in Dyslexia*, edited by John Money, Associate Professor of Medical Psychology and Paediatrics at Johns Hopkins University (Baltimore, The Johns Hopkins Press, 1962) : " It is not at all rare in psychological medicine, nor in other branches of medicine, that a disease should have no unique identifiying sign, that uniqueness being in the pattern of signs that appear in contiguity Out of context, each sign might be encountered in other diseases, or, in different intensities, in the healthy. Specific dyslexia is no exception in this respect."

A BRIEF COMMENT

ON

THE EFFECT OF SOCIO-ECONOMIC STATUS ON THE SCHOLASTIC ACHIEVEMENT OF SIXTH-GRADE CHILDREN, by ROBERT L. CURRY*

By R. R. DALE

(*University College, Swansea*)

The importance of Curry's admirable article increases the danger that his one faulty conclusion may be repeated by research workers without a careful appraisal of the data from which his finding is drawn. His first conclusion (page 48) is that " Socio-economic status seems to have no effect upon the scholastic achievement† of sixth-grade students when the students have high intellectual ability " Unfortunately his data do not warrant such an assumption. Even within his high intellectual ability group, the same trends are apparent as within his low ability group, the upper socio-economic class gaining superior scores in all cases. The lack of statistical significance for these superior scores may be due to the smallness of the samples which the experimenter has selected. We cannot say that there seems to be ' no effect ' when the effect merely fails to reach statistical significance. All we can say is that the effect may have occurred by chance ; we certainly cannot say in this case that it has probably occurred by chance, and to say that it has certainly occurred by chance—which is the equivalent of saying that in this case socio-economic status has no effect—is statistically unthinkable.

There is also a second fact which deserves consideration In work which produces a patterned result there is a tendency for us to ignore the importance of the pattern when we are considering questions of statistical significance. In Curry's results there are in effect eighteen differences between socio-economic groups to be considered ; of these sixteen are in the ' expected ' direction, and this in itself has its statistical importance. If we take the extreme socio-economic groups, comparing upper with lower, there are nine differences in mean mark to be considered, and all nine are in the ' expected ' direction The ' t ' ratios, therefore, have a cumulative effect, and although when we compare the attainment levels of different social classes the individual ' t ' ratios are lower in the high intellectual ability groups, we might well conclude that " the trend is continued but lacks statistical significance ; such significance might be obtained in future experiments by increasing the size of the sample." We need to bear in mind, also, that although we have by consent drawn an artificial line between ' significant ' and ' non-significant,' there is in actuality no such barrier, but a gradual transition.

---

* This *Journal*, May, 1962, p. 46.

† In Reading, Arithmetic and Language.

RESEARCH NOTE

THE RELATIVE PERSONALITY ADJUSTMENT OF SEVERELY DEAF AND PARTIALLY DEAF CHILDREN

By L. R. BOWYER, A. MARSHALL and K. WEDDELL.

The policy, consequent on improved electronic aids, of placing partially deaf children in ordinary schools, has raised a question of their need for special help with social and emotional problems. One hypothesis is that a partially deaf person may have more of a personality problem than one who is severely deaf.

A pilot study was carried out in Bristol between May and September, 1962, to try out ways of investigating this hypothesis and to see if there was evidence to warrant more time-consuming research. Three groups were studied, each with ten children, respectively severely deaf, partially deaf, and of normal hearing. The age range was restricted to 8–10 years. The groups were matched by paired (or triple) comparison for age, sex and intelligence. Ranges of I.Q., rather than single I.Q. figures were matched, corresponding to the Stanford-Binet and Wechsler levels of Superior, Bright, Good Average, Average, Low Average, and Dull or E.S.N. For this purpose a judgment was made based on recorded test scores, teachers' opinions, and the results of a Goodenough Draw-a-Man test. The severely deaf children (Grade III deafness) were all from one Special School for the Deaf. The partially deaf children (Grade II deafness) were either members of a deaf unit in an ordinary primary school or were individuals in classes of children with normal hearing.

THE CRITERION OF ADJUSTMENT.

The criterion of adjustment devised for this study was one of ' adequacy of fantasy,' assessed from a Lowenfeld ' World ' and a Lowenfeld Mosaic pattern, with consideration also of qualitative features in the Draw-a-Man test. This criterion was adopted for two reasons. First, the Lowenfeld techniques do not need words either in the instructions or in response, thus equating the communication problem for all groups.

Secondly, a recent American study of deaf children, (Brill, 1960) which used the criterion of how much each child was liked, showed a trend in favour of the severely deaf children, but it was not statistically significant. The authors think that this may have been because they omitted to separate those with deaf parents from those with hearing parents. In addition, however, the criterion of adjustment may have been too simple or one-sided. The present attempt to gain an idea of the inner experience of the children with the aid of projective techniques is aimed at balancing the picture from the other side. In a larger piece of research it would be important to use both ideas. In our pilot study use was made of the Vineland Social Maturity Scale, and we are grateful to Mr. Geoffrey Herbert for helping with several of these and with Goodenough and Mosaic tests. Attention was also given to teachers' knowledge of the children's homes. Only the ' adequacy of fantasy ' results will be reported here.

THE SCORING SYSTEM.

The ' World ' and the Mosaic patterns were scored in the following way, with one point for each aspect :

SAND TRAY			MOSAICS		
Recognisable Theme	Adequate Content (more than 4 categories in relationship)	Freedom from adverse Clinical signs	Good Gestalt	Freedom from adverse Clinical signs	Total possible Score
+	+	+	+	+	5

The validity of the World ' Clinical signs ' (empty, unpeopled, rigid, over-fenced, aggressive) had been established in previous research (Buhler, 1951 ; Bowyer, 1956) and also for the Mosaics (empty centre, frame design, etc , Lowenfeld, 1952).

RESULTS.

The average scores for the severely deaf, normal, and partially deaf children were 4·1, 3·4 and 2·85. With the Sign test of statistical significance (Siegel, *Non-parametric Statistics*, p.68), using the hearing group as a control or norm, the severely deaf children in this study are above average in adjustment (N 10, x=3, p= ·05). However, the partially deaf children are not below average to a statistically significant degree.

The results of comparing the two deaf groups, in matched pairs, equated for sex, age, and intelligence, gives a difference in favour of the severely deaf which is not statistically significant, although it appears noticeable to the naked eye, so to speak. A colleague has suggested that a more powerful statistic, such as analysis of variance, might bring out any significance that may be there.

It is possible to obtain a significant difference between the two deaf groups with the Sign test, if one abandons paired comparison, regarding it as having served its purpose in equating the groups. With the scores arranged randomly, p= ·002 can be obtained. By pairing (*a*) for intelligence regardless of age or (*b*) for age regardless of intelligence, then probabilities of ·06 and ·05 are obtained. Perhaps this means that mental age is the most important consideration for adjustment in physical handicap, since combining age and intelligence reduces the probability of difference in adjustment, thus :

Severely Deaf		*Partially Deaf*	
J B.	5+	V.W.	3½
S.B	5+	V.G.	2
S.H	3 –	S.R.	4
J.A.	2½ –	C.R	4
O.C.	5+	G.B.	2
M.H.	5=	S.S.	5
R.G.	4+	M.T.	1
R.R	5+	P.E.	2½
J.F.	3 –	J.R.	3½
A.R.	3½+	R.M	1

N 9, x=3, p= ·254

In the above table, the children are matched for age, sex, and intelligence. These various results seem to offer some encouragement to continue research on relative differences between the severely and the partially deaf.

THE NEXT STAGE

Two studies are planned. One is to look at the playground behaviour in those schools with deaf units, to find out how many social contacts take place between partially deaf and hearing children. The other is to proceed with a larger sample of children in a school which teaches both severely deaf and partially deaf children For it may be that the partially deaf child is less well adjusted only when he is educated as a member of a minority group in an ordinary school for hearing children, i e., adjustment may be a function of the Special School with its entire staff, buildings, equipment and grounds devoted to helping the handicapped child. The alternative hypothesis would be that even in this favourable environment the partially deaf child by the nature of his handicap, is a ' marginal man ' who fits in neither with deaf nor hearing children

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

A Survey and Discussion of the Use of Reports to Parents in Scottish and Other Schools.

By F. J. R. HONIGMANN

(*Summary of thesis presented in part fulfilment of requirements for the degree of Ed.B. in the University of Aberdeen in* 1962)

AIM.

To review the practices in the issue of reports to parents; to compare the practice in Scotland with that in selected parts of England, Wales and the United States; and to consider the ideal form of such a report.

DATA.

Thirty-five education authorities in Scotland, twenty-two in England, nine in Wales and six in U.S.A. (total seventy-two) were asked to submit specimens of reports to parents. Twenty-nine from Scotland replied, providing a reasonably representative sample for Scotland. Replies from Wales and U.S.A. (large cities only) were also representative. In England, the responsibility for issue of reports seems generally to be left to individual schools; requests were sent to a sample of English schools but less than half sent specimen reports.

METHOD.

Reports from secondary schools were selected for examination. Points considered were: frequency of issue; format; nature and quantity of information given (examination marks, standards of comparison, ratings, space provided for remarks, aspects of performance assessed).

RESULTS.

Scottish reports tend to be restricted to an impersonal list of examination marks on formal subjects with very little provision for verbal comments or assessment of personality development, in contrast with English, Welsh and American reports.

CONCLUSIONS.

Developments in educational practice have not been reflected in the form of reports to parents in Scotland. Alternative forms which communicate information more effectively and aid parent-teacher collaboration are suggested.

An Evaluation of Readability Formulae in assessing Age-Placement of Text Books.

By MARGARET TAYLOR

(*Summary of thesis presented in part fulfilment of requirements for the degree of Ed.B. in the University of Aberdeen in* 1962)

AIM.

To study the validity of three American readability formulae, Lorge, Dale-Chall and Devereux, as applied to the age-placement of twelve British school reading books graded by publishers as appropriate for children aged 9 to 10 years.

METHOD.

The criterion used for evaluating the formulae was the combined judgement of forty-five experienced teachers. An order of difficulty among the twelve reading books was calculated from the teachers' rankings, and the order obtained from the application of each of the formulae was compared with this criterion. The teachers also assessed other factors which contribute to the age-placement of the books but which are not measured by the formulae, such as format, appropriateness of content, exercises and illustrations.

RESULTS.

The teachers' ratings showed an average intercorrelation of over +0·7 and, their combined judgment was, therefore, accepted as a reasonably consistent criterion. The Lorge and the Dale-Chell formulae showed moderately high correlation with this criterion (+0·7 and +0·6 on rank-order), while the Devereux formula gave a correlation of only +0·5. All three formulae correlated highly with each other (inter-correlations between +0·8 and +0·9). Age placements by the formulae tended to be too high by comparison with teachers' opinions. There was a correlation of +0·3 between the teachers' order of difficulty and their ranking of the merit of the books as school reading texts. Additional comments from teachers and ratings of other aspects were found to add relevant information beyond that given by the formulae.

CONCLUSION.

The value of two of the readability formulae was established, but only within the limits of what these formulae aim to measure. Readability, however, is shown to be only one of several factors relevant to the appropriate age-placement of these school texts.

A Gaelic Translation of the Wechsler Intelligence Scale for Children.

BY MARY MORRISON

(Summary of thesis presented in part fulfilment of requirements for the degree of Ed.B. in the University of Aberdeen in 1961)

AIM.

To test the validity of a new Gaelic translation of WISC, with reference to the study of the effect of bilingualism on intelligence.

SUBJECTS

All Gaelic speaking children in Lewis, born in April to September 1948 (inclusive), aged 11 : 6 to 12 : 0 at the time of testing. Of 130 such children, 128 were tested.

TESTS.

In addition to WISC (Verbal and Performance Scales), two Moray House Verbal Reasoning Tests, Raven's Progressive Matrices, teachers' estimates of school performance, and indices of the occupational status of the homes and of the extent of Gaelic in the linguistic background of the homes.

RESULTS.

The main findings were :

(i) The Gaelic translation of WISC agreed with the Moray House tests and Matrices in showing an average score of 92 or 93 in the group.

(ii) Gaelic children were at a disadvantage on performance tests, and on strictly timed tests.

(iii) The form of WISC sub-tests recommended for use with Gaelic-speaking children are :

> Verbal—Comprehension, Similarities and Vocabulary (all in Gaelic), Arithmetic and Digit Span (both in English).
>
> Performance—Picture Completion, Block Design, Object Assembly, Coding and Mazes (all with Gaelic instructions).

The Effects of Treatment in a Special Class in a Junior School on Retardation and Maladjustment with Special Reference to the Concept of Self

BY DAPHNE C. IVES

(*Abstract of thesis submitted for the degree of M Ed in the University of Manchester, December,* 1961)

AIM OF THE INVESTIGATION

This research was undertaken in order to assess the value, to eleven children of junior school age, of a year spent in a special class The class was attached to a primary school in a Lancashire cotton town which accommodated 350 children The children had been placed in the class for one or more of the following reasons they were borderline cases for placement in a school for E.S N. children ; or, they were of low-average intelligence and severely retarded in one or more of the basic subjects ; or, because of some emotional disturbance, they had been unable to adjust themselves to a normal classroom situation.

Special emphasis was put upon changes in the self-concept of these subjects.

METHOD

(1) Measurements and evaluations made at the beginning and at the end of the period included Sander's tests of Social Under-Evaluation (SUV) and Non-Social Tendency (NST),[1] Raven's Projection Test,[2] the Bristol Social-Adjustment Guide, forms 1 and 2,[3] together with tests of persistence and level of aspiration designed by the present writer. In addition, a series of controlled observations were made by the writer, using a check list during free-play periods. Items which were checked were subsequently grouped under four separate headings · hostility, social interaction, withdrawal and constructiveness All of the data so collected were analysed by the Wilcoxon Matched-Pairs Signed Ranks test[4] to establish the direction and significance of changes over the year.

(2) Additional information was provided by (*a*) questionnaires which were completed at the beginning and end of the year on the basis of separate interviews with the children themselves and with their parents or guardians , (*b*) questionnaires administered to sixty-six children who attended the school to which the special class was attached concerning their attitudes to the children in this class ; (*c*) questionnaires completed both by the special class children and by a control group taken from the remainder of the school regarding their attitudes towards previous teachers ; and (*d*) four sociometric tests, covering the period

Although none of the data described in the previous paragraph lent itself readily to statistical analysis, the information was found valuable in forming a qualitative picture of the effects of placement in a special class of this sort.

(3) Educational progress was measured by standardized tests of attainment administered at the beginning and end of the year. In spite of what appeared to be marked progress made by the majority of the subjects, however, the results did not lend themselves to statistical comparison owing to the fact that many of the children had failed to make any score on entry.

RESULTS.

(1) Significant changes were found to have occurred in the areas of adjustment measured by a number of the tests used. Social under-valuations had decreased ($P < \cdot 01$) as also had non-social tendencies ($P < \cdot 025$). Although the total number of adverse pointers listed in the Bristol Guides showed no significant change, there was a significant reduction in items classified U/W (unforthcomingness—withdrawal) and K (Knavery or psychopathic hostility to the institutions of society and its representatives as they impinge on the child), ($P < \cdot 025$ in both cases). The subjects displayed a more realistic approach towards future performance, as shown by a decrease in the final discrepancy scores calculated from the test of level of aspiration ($P < \cdot 025$). Observed incidence of hostility showed a marked decline during the second half of the period ($P < \cdot 01$) and constructiveness showed a corresponding increase ($P < \cdot 01$).

No significant changes appeared in ' conformity ' as shown by Raven's Projection Test,[2] in ' persistence,' or in the observed frequency of interaction or incidents classified as ' withdrawn ' during the controlled observations.

(2) The qualitative results of the experiment may be briefly listed as follows :

(*a*) The responses to the questionnaires given to the subjects indicated that they all had a poor self-view in relation to school work at the beginning of the year Ten of the subjects showed an improvement in this respect by the end of the year.

(*b*) The results obtained from the initial questionnaire given to a parent or guardian of each of eleven subjects, showed that ten of the eleven adults were prepared to co-operate with the special class teacher and were grateful that special help was being provided for their children. At the end of the year, eight of the parents were very satisfied with the social and academic improvements made by their children. One (a guardian) was not interviewed again and the other two had certain reservations about the class, but were not wholly dissatisfied.

(*c*) The responses to a questionnaire given to sixty-six children in other classes within the school indicated that some of them regarded the special class children as being inferior to themselves. Remarks to this effect were made in the playground which could have had an adverse effect on the self-concepts of the subjects.

(*d*) When the subjects were questioned about their attitudes to teachers, the incidence of self-referential statements was higher among the special class children in both the ' Teachers Liked ' and the ' Teachers Disliked ' categories than among the ' normal ' children. This indicates that the self-forming influence of teachers is greater with extremely unsuccessful children.

(*e*) The sociometric inquiries showed that of three children who were socially isolated within this group at the beginning of the year, one had become satisfactorily integrated by the end of the year.

(*f*) Attainment tests indicated that ten of the eleven subjects had made sufficient academic progress by the end of the year to enable them to feel that they were beginning to acquire the basic skills.

CONCLUSIONS.

It should be emphasised that because of the small sample and the limitations of the experimental situation, the following conclusions need to be regarded as tentative.

1.—During the year of the experiment the subjects who had been designated ' withdrawn ' (criterion : ten or more U/W pointers on the Bristol Guides) became significantly less so. It can be tentatively concluded, therefore, that relative success, or a permissive classroom climate, or both, help such recessive children to make a more adequate social-adjustment.

2.—There was a decrease in hostile behaviour and an increase in the number of constructive interactions observed in the second half of the year. This suggests that providing opportunities for free interactions to children who have failed to make a good social-adjustment is beneficial, in so far as it helps them to function more adequately in a group.

3.—After a year in the special class, the subjects tended to make their predictions on the level of aspiration test more in accordance with reality. As the level of aspiration situation is directly related to feelings of success and failure (a high discrepancy score indicating feelings of inadequacy) it is concluded that the decrease in scores is indicative of an improvement in the self-concepts of the subjects which could be attributed to one or more of the following reasons: (*a*) most of the subjects had begun to master the basic skills of reading and arithmetic for the first time in their lives; (*b*) they were feeling more socially competent because of the opportunities provided for social rehabilitation; (*c*) significant people in their lives were encouraged to provide self-enhancing experiences for them.

4.—A close liaison with the parents was considered to be an important factor in helping the children, both academically and socially.

In the judgment of the writer it would be desirable for such special classes as this to be attached to schools which are oriented towards individual progress rather than towards competition in the form of examination positions and streaming by ability. Competition in this form leads to a certain amount of stereotyping which is not conducive to healthy self-evaluation on the part of the children who are unsuccessful within the system. It seems particularly important that children with problems of academic failure and social maladjustment should be placed with teachers who are temperamentally able to accept them and deal with them in a constructive manner.

REFERENCES.

[1] SANDERS, C. (1948). Insecurity and social maladjustment in children. *Brit. J. Educ. Psychol.*, Vol. XVIII, Part III, pp 148-155.

[2] RAVEN, J. C. (1951) *Controlled Projection for Children.* London: H. K. Lewis and Co., second edition.

[3] STOTT, D H. and SYKES, E G. (1956). *Bristol Social-Adjustment Guides Nos.* 1 *and* 2, U.L P.

[4] SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill Book Co.

The Opinions of Young Workers on their own Education in a Secondary Modern School.

BY FRANCIS S. JACKSON

(*Abstract of thesis submitted for the degree of M.Ed. in the University of Manchester, April,* 1962.)

PURPOSE OF THE INQUIRY.

The purpose of the inquiry was to obtain at first hand the adult opinions on education of a group of ex-pupils of a secondary modern school and to examine these from the point of view of (*a*) any relevance they might have to the organization and administration of such a school, and (*b*) their bearing upon how the school stands in the esteem of its own ex-pupils, the parents of tomorrow.

THE SAMPLE.

This consisted of 142 young men drawn from four consecutive one-year age-groups. They had left the same secondary modern school at the age of 15 and were, at the time of the inquiry, between 20 and 24 years of age They represented 83 per cent. of the school's total output of 15-year-old leavers over the four-year period in question, and included a wide range of occupations ranging from undergraduate to labourer. A simple statistical check revealed no significant differences in age and ability-grouping between the 83 per cent. included in the sample and the 17 per cent. omitted.

THE SCHOOL.

The school was a small secondary modern boys' school serving a sea-side town and the adjoining semi-rural area. Its organization and size (two form entry, with a single fourth form) was fairly stable during the period under review. The curriculum incorporated no particularly distinctive features and no unusual teaching methods were in use.

PROCEDURE, AND SCOPE OF THE INQUIRY.

The majority of the respondents (115 out of 142) were interviewed personally by the investigator in their own homes, usually in the evening. Those who, for various reasons, could not be reached in this way completed a questionnaire which covered the same topics as were dealt with in each interview. These were :

(*a*) Respondent's present occupation.
(*b*) Record of further education (if any).
(*c*) Opinions as to the adequacy or otherwise of the school as a preparation for employment.
(*d*) The respondent's views on the relative values of various school subjects.
(*e*) Opinions and ideas on vocational guidance.
(*f*) Free comment or criticism on any seemingly relevant matters
(*g*) An over-all assessment of each subject's satisfaction with his own secondary education on a seven-point rating scale.

SUMMARY OF CONCLUSIONS.

1.—The majority of the young men in this four-year sample of ex-pupils were favourably disposed towards their own secondary modern school There was a slight tendency for those in the upper half of the ability range to be more critical (constructively) than their fellows Extreme dissatisfaction was rare.

2.—There was a distinct association between ability group and strength of demand for higher academic standards. A higher standard of mathematics was particularly requested The requirements of further education appear to be largely responsible for this.

(3) Almost 72 per cent. of the entire sample (over 94 per cent. of the ablest quarter) had engaged in some form of further education. This had led, in two cases, to the university, and in many others to diploma and certificate courses of various kinds. Such achievements as these, especially in ex-pupils of so small an establishment, need to be remembered by those who would dismiss the secondary modern school as an academic ' dead end '

4.—Suggested additions to the curriculum (as it then existed) were metalwork, a modern language, and algebra. Requests for metalwork were numerous and appeared to bear no relationship to level of general ability Demands for algebra, on the other hand, showed a distinct association with level of ability.

5.—Many of the ' second best ' performers felt they had not received their full share of attention, particularly in the basic subjects. They felt they should have been pushed harder, encouraged, and allowed to try some of the work which was evidently considered beyond them.

6.—Music and history came in for heavy criticism as serving no worth-while purpose in the modern secondery school. From the less severe critics came suggestions that music might be made optional and that a new approach to history might be made via civics and current affairs

7.—Mathematics, English, practical subjects, science, and geography were rated by *all* ability groups as the five most generally desirable subjects. From a purely vocational point of view, mathematics, woodwork, and technical drawing were rated higher than any other activities.

8.—Age, as reflected in the Christmas, Easter and Midsummer leaving groups, appeared to be significantly associated with level of ability, the upper ability groups containing a disproportionately high number of the Christmas leavers. Since this phenomenon has been noted elsewhere, its possible effects deserve further study

9.—Approximately half the respondents declared themselves satisfied with existing vocational guidance arrangements, but some of the more thoughtful suggestions made pointed to the value of adequate school records and to the possibility of the detection of specific abilities as they emerged.

10.—The level to which these young men have risen in their chosen occupations, their general record of stability in employment, and the level to which they have carried their education since leaving the secondary modern school suggest that for the majority of the 142 members of this sample their education at such a school has proved to be a not inadequate preparation for entry into the first phase of adult life.

BOOK REVIEWS

ANSTEY, E. (1961). *Staff Reporting and Staff Development.* London : Royal Institute of Public Administration in collaboration with George Allen and Unwin, Ltd., pp. 96, 12s. 6d.

Applied psychologists who have read publications by Edgar Anstey would expect this book to be meaty, concise and practically helpful. They will not be disappointed. A quarter of the book is made up of appendices showing exact copies of staff reporting forms in current use. In the remaining seventy odd pages Anstey is able to compress, in a highly readable form, balanced accounts of the objectives and methods of staff assessment, of confidential and open systems of reporting, the pros and cons of Promotion Boards and some very worthwhile advice on the design of report forms ; he gives us the benefit of his experience in training in reporting and describes three kinds of staff appraisal interviews. A section which particularly impressed me was the five-page critical account of the forced-choice technique. I also liked the examples he gave of the ways in which clever reporting officers could subtly twist the reports so as to secure the promotion of people they liked or the transfer of people they disliked. Obviously, Edgar Anstey has worked with very intelligent people.

On the face of it, this book is a valuable one for Establishment Officers in County Halls and Ministries but not one for educationists. However, those concerned with the grading and assessment of children, of students and of teachers and those who work with follow-up and validation in any field . . . all these will find *Staff Reporting and Staff Development* a very useful handbook. DENIS MCMAHON.

BEREDAY, G. Z. F., and LAUWERYS, J. A. (1962). *The Year Book of Education—The Gifted Child.* Evans Bros., in association with The University of London Institute of Education and Teachers' College, Columbia University, New York, pp. x+541, 63s.

Readers of this *Journal* will probably find this volume of the Year Book one of the most valuable of the series for it deals with theoretical and practical problems of giftedness to which society is urgently demanding answers of psychologists. The three major sections are headed *Psychological Theories, Detection of Ability and Selection for Educational Purposes* and *Increasing the Pool of Talent.* These are preceded by an introduction by Sir Cyril Burt which must surely be one of the most distinguished pieces of writing in educational psychology in recent years. Here in under sixty pages he broadly surveys the whole field and so orientates the reader for the detail which follows and provides him with a critical evaluation in advance of much of what appears in the rest of the book.

Section One deals with theories as to the nature of giftedness and the educational practices likely to stem from them. Of particular interest are Ruth Strang's introduction, Guilford's *Parameters and Categories of Talent* and Torrance's *Measurement and Development of Creative Thinking Abilities.* The methods of selection of gifted children in a wide range of countries which are described in the Second Section are on the whole retrospective studies reflecting existing techniques which have grown up in the countries of their origin with little reference to advancing psychological knowledge. Most are based upon examinations and frequently include tests of academic aptitude. Some appear unsophisticated to British eyes. There are, however, some valuable accounts of researches designed to assess the efficacy of procedures, as, for example, in Passow's survey and Husén's work in Sweden into the age at which selection should take place. Contrary to many beliefs Husén concludes that early selection of the academically gifted does not benefit either those selected or those left behind.

The third section concerns attempts of modern technological communities to meet their insatiable demands for people of high talents, and of rapidly developing countries to increase their available human resources as quickly as possible. The proposed solutions are the familiar ones of recognition followed by appropriate educational environments designed to encourage maximum development commensurate with individual genetic limitations. There is less than one might have hoped of discussion of the social problems satirised by Michael Young which inevitably arise as segregation of the most able becomes more efficient, though Anderson does touch on them in his chapter on *Dilemmas of Talent Centred Educational Programmes.*

The levels of writing are inevitably uneven and few but a reviewer will have the fortitude to read the whole, but there is no doubt of the book's value for it gathers together a great range of thinking on the topic of giftedness and provides a very complete entry into the many aspects of the field. L. B. BIRCH.

CLEUGH, M. F. (1962). *Educating Older People.* Tavistock Publications, pp. 185, 28s.

This interesting book is an attempt to generalise the experience of the author in tutoring students for the diploma in teaching backward children working at the Institute of Education, University of London The author makes considerable use of what these students have said themselves in reply to queries about their experience of the course. The book has all the virtues and defects which arise from the use of such a method : flashes of insight occasionally light up the general picture of banality and do-it-yourself psychoanalytic rumination which constitutes the second half of the book. The treatment is very uneven since Dr. Cleugh seems to work at two different levels, as a rational analyst of that experience which she shares with trained teachers (the most fruitful part of her communication) and as an amateur case-worker who works on the most superficial and subjective level of analysis of ordinary human contact. In the days of my youth I became familiar with the kind of content which is now being widely publicised as the latest advances of psychological insight—except that in those days it was described, perhaps unkindly, as " stair-heed gossip." I quote :

" *January* 29.

Tutor : Are you keeping your head above water ?

Mrs. S. : Yes, much better this term.

Tutor : Good. Take it easy.

Mrs. S. : It's only *just* above water, but it is above. (*Seems steadier and more relaxed, in spite of brick on 17th.*)." p. 129.

The first hundred pages deal in a systematic way with the problems of educating older people, under the chapter-heads of *Getting things done, lecturing, discussion, supervision of practical work, examinations.* This part of the book might be regarded as a handbook for adult tutors . it is helpful, informed and systematic Although it is on an *ad hominem* level with very little reference to the scientific studies which have been made of the adult learner, there is nothing which one would fault as contrary to one's experience or conclusions about the adult.

The second half seems to have been the result of some kind of pressure to produce a " proper book " in the usual Tavistock format and style rather than a mere brochure of " suggestions for the use of tutors." Dr. Cleugh expresses the view that although the tutor is not a professional therapist he can (and *ought* to ?) serve a therapeutic purpose. It is the attempt to become self-conscious and centrally concerned about this function with which one would quarrel.

However, there have been so few books published in this country dealing with the adult student that one must compliment the enterprise of the author and the publisher, and hope that the book will be widely read since it will certainly be of value to many different kinds of reader. Although not the most economical in terms of time, one occasionally is illuminated in listening even to " stair-heed gossip."

JOHN MCLEISH.

GORDON, I. J. (1962). *Human Development.* Harper, pp. 400, 46s.

This book traces the development of the child from conception to manhood, preserves a good balance between the forces of heredity and environment, and successfully combines the physical, psychological and sociological aspects. The adjective which best describes the first two-thirds of the book is ' sound,' and this is by no means damning with faint praise Two late chapters, on " Exceptional Development " and " The Continuing search for Identity " are very good. In general, both effectiveness and interest tend to increase as one reads on.

The author endeavours to study the developing child as a whole while also examining the forces bearing on him at certain times in his life, and seeks to adopt a ' unifying theoretical position.' The development of the self ' is a transactional process between the developing organism and its changing social environment.' He tries, to combine an " external " approach of looking at the child's home, his peer group, his school and his society, with the " internal " approach of looking at how he perceives these various situational forces, and how he perceives his own body and his own growth.' The theoretical account which is constructed on this basis is well supported by concise but informative statements of recent research. Though there are some references to British work, there could profitably have been more.

Though the book is particularly sound there are a few statements which need careful examination—though one indeed has seen them elsewhere. Gordon adumbrates that " The direction of (man's) behaviour is always, in normal development, towards increased complexity." The intention here is good, but one wonders about the word ' always.' When man, for example, reduces a chaos of isolated facts into order by classification is he not simplifying and consolidating the position, maybe preparing for further complexity but not increasing it for the moment ? Similarly, we read " the increase in complexity and organisation of the person continues throughout his life." Does this statement include old age ? Much more important, however, is the widely accepted theory of ' equilibrium ' which Gordon states once again. The human organism " tries to maintain a steady state, i.e., it attempts to maintain itself and to balance the forces within itself " While there is a central truth here the words do not describe exactly what is happening, and Gordon (and many other writers) contradict themselves when they maintain that the organism "at the same time develops towards a new steady state." The reviewer's position is that the developing child must contain, if healthy, a tendency towards greater maturity, towards development, therefore *ipso facto* it cannot be in a ' steady state.' Physiologically, it is obvious that the child contains within himself the tendency to change, and this is also true psychologically. This often repeated concept of ' equilibrium ' needs a much deeper examination and a much more thoughtful comment.

R. R. DALE.

HADFIELD, J. A. (1961). *Childhood and Adolescence.* Pelican Books, pp. 286, 5s.

There are few books of child psychology, covering development from infancy to adolescence, which can be recommended confidently to parents who lack specialist knowledge of psychology. Dr. Hadfield has provided such a book by writing for " ordinary intelligent parents of ordinary intelligent children." He has drawn on his experience as a psychiatrist to give many examples of children's behaviour and difficulties, and to make recommendations to parents about the treatment of children. These are founded on his belief that the fundamental need for normal development is a sense of security. It follows that he gives little emphasis to sexual problems as primary causes of neurotic symptoms ; he considers that mental health in childhood depends on parental love, on discipline which combines kindness with firmness, and on freedom from the inhibiting effects of unnecessary restrictions or ' over-protection.' Some readers may feel that genetic constitution and social environment also play an important part but, in so far as Dr. Hadfield considers the latter, he finds that a sense of inferiority and competitiveness play minor roles in the development of most children.

Following an introductory discussion of parents' problems in bringing up children, chapters two to four describe the stages of maturation, phases in early development and the organisation of personality during childhood. In chapter five emotional development in adolescence is described at some length. More serious neurotic disorders are dealt with briefly in chapter six and, in a last chapter, several ' golden rules for parenthood ' are suggested.

Dr. Hadfield stresses the importance of allowing children to live fully through each phase of development, expressing each impulse before it is disciplined and directed. In disciplining children, he recommends the use of encouragement with a minimum of blame ; but he considers that punishment of wrong-doing should be inevitable, with the proviso that the child's motives should first be understood.

The long chapter on adolescence is likely to be most appreciated by parents ; for whereas writers in some popular journals deal knowledgeably with the difficulties of young children none deals so soundly with the problems of adolescence. The four stages of sexual development outlined here may be a new conception to readers already having some knowledge of psychology.

This is a well-written book, in a pleasant and persuasive style with a fund of illustrative anecdotes ; the views expressed are likely to impress most parents as eminently sensible. It should do much to help them, as well as stuents and teachers, to understand the development of children and to alleviate their difficulties

RUTH BEARD.

HAWKES, GLENN R., and PEASE, DEMARIS (1962). *Behavior and Development from* 5 *to* 12. New York : Harper and Bros., pp. ix+375, 46s.

This is rather a charming book which one would wish to have on one's shelves (rather than order for the Library). It has an attractive format, clear print and good photographs. The content is remarkably free from distortion, considering the brevity of treatment of a wide variety of topics These are presented in fifteen chapters, each with headed sections and sub-divisions. The integrity of approach, and the deft, direct style, impose their own unity, and can stem only from being thoroughly at home with the subject-matter, and having masterly powers of selection. The book is at once easy to read and motivating to more advanced study, for which references are given at the end of chapters. The work of British writers is included (e.g., for Ch. 9, " growth and function of intelligence " mention is made of research by C. Burt, M. Dunsdon and J. Fraser-Roberts, and G. Seth).

The authors state that their book is about children in their natural habitat—their homes, their neighbourhood and their schools. The reader who expects to have a clearly separated view of each year from 5 to 12 will not find this, except for a very condensed " summary of age characteristics " (agreeing with Gesell's accounts) in Ch 14 (pp. 315-330). The emphasis is more on the 5–12 ' school-age,' as a whole, except for incidental differentiation.

The book is ideal for groups of the weekly tutorial or seminar kind, either in training colleges, adult evening classes, or in university introductory courses in child development with a social psychology orientation. At the end of most chapters is a list of " Study Problems." These are thought-provoking practical exercises not too onerous to be carried out with enjoyment in limited time. There is a chapter (13) which gives details and principles for observing and recording children's behaviour. Appendix A and Appendix B list, respectively, films and selected pamphlets, each with a descriptive sentence. The last chapter is " Adolescence, a look ahead " with suggested readings, which would give a study group the basis for a follow-up course.

It is assumed throughout that the parents and teachers for whom the book is mainly written will wish to use the knowledge provided. There is also a specific chapter (12) on guidance, with a word of caution in the preface : " The understanding parent or teacher will refrain from the use of specific techniques unless he has a philosophical understanding of the child and the demands being made upon him."

A lively feature is the introduction of quoted extracts, sometimes from children themselves, e.g., exchanging jokes and riddles (p. 227), or an 11-year-old writing about teachers (p. 53) : " There are some teachers who have just gotten tired of children because of the bothersome things they do, and do not like any of them any more. This kind do not help you learn very much . . " The authors of this book are clearly not in this category. " Being around children," they say (p. viii) " may make the hair gray, but it keeps the heart young " It is pleasant to be able to recommend a book of high academic standards written with enthusiasm in a manner which meets the needs of the less academic and of those students and professional workers who would benefit from a moderately short, gay but not superficial text for this relatively little studied 5–12 age. L. R. BOWYER.

MANNHEIM, KARL, and STEWART, W. A. C. (1962). *An Introduction to the Sociology of Education.* London : Routledge and Kegan Paul, pp. xvii+ 187, 21s.

This work is chiefly of historical interest. While it may be the last of Mannheim, no one would claim that it was the best. Its fifteen chapters, arranged in five sections, are uneven in value and loosely strung together. It purports in its title and Introduction to constitute an introduction to the sociology of education. It does not. No beginner who read this book would, in fact, have been introduced to the major pre-occupations, problems and achievements, still less the methods, which characterize educational sociology in the mid-twentieth century

One-third of the book is devoted to psychological topics : resumés of the more obvious and unsophisticated theories of learning and personality development. There is some advance beyond associationist theory : the laws of recency, frequency and effect are rehearsed, but attention is given to the social context in which learning takes place, to the assumption of roles within social institutions. This is both too little and too much. It is doubtful whether the beginner would make much of a passing reference to Allport's theory of ' functional autonomy ' ; for others it is really rather old hat. The chapters on social influences on personality development draw heavily on G. H. Mead ; they are written with only a bare hint of the contribution which the cross-cultural studies of social anthropologists have made to this field of inquiry.

The first third of the book and the last deal with matters which are more properly sociological. It is interesting, however, that in a chapter which usefully discusses teaching, training, and instruction, the temptation to explore the sociological implications of these distinctions is firmly resisted. The sociological chapters contain much that is platitudinous and trite. An ambitious survey of the ideals of the Graeco-Roman, Hebrew, Medieval, Renaissance, and post-Renaissance worlds illustrates the point that different ages modify the ideals (for example of ' the gentleman ') which they inherit, and that ' current attitudes have historical antecedents.'

Nowhere in the book are we introduced to the sociology of education as an experimental science concerned with the rigorous empirical testing of hypotheses. In the final chapter, it is true, a number of suggestions are made for research : a ' survey of the existing educational institutions in this country and their interrelations . . . ' ; the factors in modern life (work and leisure) which make for the disintegration of the personality ; the proper recruitment and role of élites if the cultural decline of our society is to be halted ; ' the balance between the contributions of the home, church, school, factory, and public life towards moulding individual character.' There is little apparent awareness of the formidable technical problems involved in setting up inquiries which might yield answers to these great questions. This reviewer remains unconvinced that " the right command of primary sources and of the existing literature on the subject already offers answers to a great number of them." F. MUSGROVE.

MONEY, J. (Edit.) (1962). *Reading Disability. Progress and Research Needs in Dyslexia.* Baltimore: Johns Hopkins Press, pp. x+222. No price given.

The papers published here were submitted at a symposium held in the Johns Hopkins Medical Institutions in 1961 under the auspices of the Association for the Aid of Crippled Children There are thirteen symposiasts drawn from the general fields of education, psychology and medicine. In itself, this is a reassuring sign for medicine and paedagogy have been too slow in coming together in joint studies of mutual interest

Workers in the field of severe reading disability will not be surprised that no dramatic break-through is reported or that Howes writes " the field is still at a stage where we need to find out exactly what dyslexia is as well as why it occurs." It is noteworthy, however, that in spite of the opinions given by such as Torpe and Malmquist (who were not at the Conference), that no doubt is expressed as to the existence of the syndrome of specific developmental dyslexia. It is seen as a dysfunction of written language and communication resembling the sequelae of some brain injuries.

Most of the familiar topics are dealt with and brought up to date by reference to the most recent work and there are some topics which will be unfamiliar. On the educational side there would seem to be little new to report with regard to phenomenology, or to recognition and treatment of the condition As Schiffman indicates we still await answers to the questions : What is the aetiology ? How can we make an early diagnosis ? What are the best remedial procedures ? It seems clear that answers are unlikely to arise from further large scale surveys and from correlational and factorial studies. It is pointed out by Zangwill that detailed clinical and experiemntal studies are most likely to throw light upon them. The latter half of the book is concerned with such studies though the papers are largely statements of points of view of workers in particular fields rather than the reporting of experiments.

Birch, in his paper on *Dyslexia and the Maturation of Visual Function*, insists that dyslexia is not itself a problem capable of being investigated since it may derive from a variety of aetiologies. This, he suggests, makes it necessary for any worker to decide at the outset which aspect of the problem he wishes to explore ; but, having made a decision, this does not mean that he regards other aspects as unimportant It is the main value of this book that workers who have made decisions of this kind have placed their considered opinions side by side so that each can be correlated with all others and the way forward seen a little less obscurely.

This is essential reading for all workers in this field It follows on and supplements M. D Vernon's (1957) definitive volume *Backwardness in Reading* bringing the bibliography up to 1961. L. B. BIRCH.

SCOTT, W. A., and WERTHEIMER, M. (1962). *Introduction to Psychological Research.* New York : Wiley, 60s.

The writers of this book have set out to provide help to those who are embarking upon psychological research for the first time. They use the word psychological broadly and what they have to say is relevant to all the social sciences.

There are some to whom such a book will be an anathema. They would aver that the only way to learn how to do research is to embark upon it and that without such a process the knowledge acquired by the theoretical study of research procedures is ephemeral and jejune. But this view forgets the position in which the newcomer finds himself. It ignores the fact that psychological research techniques are highly sophisticated, are not so easily acquired by contagion as in many other fields and exist in considerable variety. It does not, moreover, take cognisance of the fact that, for many new research workers, it is the existence of techniques as well as of problems which often provides orientation and a conceptual framework for their proposed studies.

The need for the guidance which such a book can provide would almost certainly be granted by those who have numbers of higher degree students passing through their hands, yet the attempts to provide it have been rather few. Of these, the present is the best which has yet appeared. It deals in a co-ordinated manner with the things one would expect, with the processes of planning and design in fair detail, with the techniques of instrumentation, verbal, observational and experimental, with sampling and with analysis, interpretation and reporting of results Throughout, the hand of experience is clearly revealed and the discussion is informed, up to date and most clearly laid out. Yet, even so, if the purpose of the book was to inform and guide the beginner it might still have been better. It tends, as do all books in this field, towards a discussion of what research is about, as if it were a discipline in its own right, rather than concerning itself with how to do research, insofar as it is ever possible to describe this. This is, of course, a simplification and generalisation which in some detail is unfair to the authors. To several of their chapters they have added supplements which deal with the ' how ' rather than the ' why ' and these chapters, and not only their supplements, are among the best in the book Yet, too much of the 400 pages of text is taken up with discussion which is excellent for the relatively sophisticated but which is inappropriate to the person seeking fairly specific guidance during his novitiate. One suspects at times that the writers wanted to present a straightforward yet informed guide book to the neophyte but were deterred from writing the whole book in this way by the fear of being thought unscholarly. If this were so, it was an unnecessary fear both from the viewpoint of the subject matter and from the standing of the authors themselves. Had they written more specifically a guide book in the way indicated, they would surely have found space to deal, *inter alia*, with questionnaires, to which so many beginners are attracted, with a treatment of non-parametric techniques in sufficient detail to give the newcomer some orientation and, surprising omission, with factor analysis in outline and with the types of study amenable to it. If space was short they could well have omitted the twenty-five pages of standard tables.

But these observations, notwithstanding, the book remains the best of its type and should certainly be in all libraries, in the hands of those sophisticated enough to skip some parts and to benefit from others and with anybody else who is not deterred by the rather high price ticket which American books seem to attract in this country

G E. R. BURROUGHS.

VENESS, THELMA (1962). *School Leavers.* London: Methuen and Co., Ltd., pp. xxv+252, 25s.

This book is concerned with young people's outlook, chiefly on their own personal future What do they expect from life in its two main spheres, family and work? In particular, and this is the problem which first stimulated the inquiry, what are their ambitions?

The study was carried out with 1,300 boys and girls in two counties, one in the West of England, and the other in the vicinity of London They were in grammar, technical and modern schools, with ages ranging from 13-plus to 17-plus, depending on leaving age. The methods employed were, to borrow a phrase from Wilkins, of the ' grape shot ' variety : questionnaires, interviews, paired comparisons and so on The underlying assumption, shared by the reviewer, is that in certain kinds of research a multitude of relatively coarse approaches is more likely to yield a rounded picture than a single approach, however thorough. The key technique on which the author relied most heavily—though without omitting frequent checks from different sources—is the ' restrospective autobiography.' The writer was asked to imagine himself at the end of his life, and review it from that perspective This technique, which was previously used by Stoetzel in his study of Japanese youth, produced a rich harvest of fascinating material.

Some aspects of ideas about marriage and the family may be cited by way of illustration. Current trends are reflected in the favourite age for marrying, between 20 and 25. Although it is not directly stated, one might infer from the quotations a remarkable absence of the romantic notion of marriage (remarkable because the prevalent stress in the mass media) · " This job I stayed at until I got married at the age of 20 I had a good husband and we never quarrelled " It is mostly rather matter-of-fact. This also applied to the references to children ; there was little to support the view that frustrated aspirations tend to be projected on to fantasy children, and scant emphasis on education as a goal. Very few mentioned separation or divorce, but a curious feature was the anticipated high mortality rate of spouses.

Attitudes about work, knowledge of jobs and job choices are extensively dealt with, and this section of the book wi ll be found most useful by those concerned with the practical problems of school leavers. The author introduced an interesting analytical categorization of job-choice orientations in terms of ' tradition-,' ' inner- ' or ' other-directed,' which deserves further development.

The findings concerned with ambitions have aroused most widespread interest and led to the vociferous airing of prejudices. A common reaction to the book was that it reveals the youth of today as dull and unenterprising ; had the results been in the opposite direction, the line would no doubt have been that modern youth is aiming too high, with inevitable later frustration. Here it will only be noted that the statistical analysis of ' ambitiousness ' items has as its outcome a scale on which technical school boys rated highest. The problem is of course,, a difficult one, and the author herself did have some doubts on this score. It is, perhaps, questionable whether any single scale can satisfactorily measure ' ambitiousness ' across the whole range from unskilled to higher professional work. Another factor, which should be kept in mind with regard to all the comparisons made between schools, is that a majority of the grammar school pupils investigated were somewhat older than the rest.

These comments are only possible because extensive and detailed tables are provided. Readers tempted by the lively style to skip perusal of the tables should resist the temptation—they contain much of the meat. Altogether the book has a great deal to offer to everybody concerned with young people, but perhaps most of all to educationists (a sample of their products) and to youth employment officers (a warning that what they are apt to regard as raw material is already partly processed.).

G. JAHODA.

VYGOTSKY, L. S. (1962). *Thought and Language.* M.I.T. Press and Wyley.

This is an English translation of a book written getting on for thirty years ago which appeared again in Russian only as recently as 1956. It shows remarkable insights into the nature of thinking, and, in its essentially functionalist approach, is an appropriate work to put before the English reader today.

After presenting his problem, which is the growth of thought as shown in the construction of concepts, Vygotsky then discusses Piaget's and Stern's views of language development. The comment upon Piaget is, of course, based only on his earlier works and the publishers include with the book a separate paper of some thirteen pages, setting out comments made by Piaget in 1962. Vygotsky is critical of the extent to which Piaget's theory leans on the concept of maturation. However, Vygotsky's own contributions form the really important parts of the book, particularly as set out in Chapters 5, 6 and 7. In Chapter 5 he describes how conceptual grouping by children grows from heaps having no objective basis, through complexes with no logical basis but possessing some of the qualities of concepts, finally to true concepts. In the next chapter, on the development of scientific concepts, he has much to say to the educationalists, in particular contrasting spontaneous with non-spontaneous concepts. The latter he exemplifies from school work in science and modern language learning, and shows how the two types of concept tend to come together in the course of learning. Finally, there is a thoughtful discussion of thought and word, in which the growth of speech and thinking are considered together and in which Piaget's views on the socialising of egoentric speech is commented upon and checked and criticised by empirical study.

In all the book is a ' must ' for workers in the field of the growth of children's thinking and language Those who know the publication by Piaget and Inhelder, *La genese des structures logiques elementaires*, will also recognise how similar is the picture presented by Vygotsky of the growth of concept formation to that set out by the Geneva School in their publication The book is, however, more than a psychological treatise. Much of it offers a method of education, and so should be read by research workers in this field. E A. PEEL.

WISEMAN, S. (Edit. 1961). *Examinations and English Education.* Manchester : Manchester University Press, pp. xx+188, 21s. net.

In 1959, the subject chosen by the School of Education in Manchester University for its annual public lectures was the problem of examination. There were, we are told, several reasons for this choice—the growing interest of secondary modern schools in examinations of an external type, the feeling that the time was ripe for taking stock of post-war changes in examinatious which affect either primary or secondary schools, and finally, the recent developments in the techniques of examining and the need for relating them to the purposes the examination seeks to fulfil. The present volume, edited by Professor Wiseman, is an expanded version of the lectures delivered on these topics, with an additional chapter from Professor Oliver dealing more especially with the recommendations of the Crowther Report.

After a short introduction, the series opens with a ' historian's view of examinations ' by Mr. Norman Morris. This is followed by lectures on examinations in relation to the grammar school, the modern school, and the primary school, respectively, by Dr.Petch (Secretary to the North Universities Joint Matriculation Board), Mr. Lucia (formerly Secretary to the Associated Examination Board), and Dr Wiseman. Dr. Wiseman also summarizes the methods and results of assessing the efficiency of examinations ; and Professor Oliver discusses ' examining general education,' and then winds up the course with a paper on the problems of education and selection.

All five contributors are conscious that both selection and examinations have their limitations and their drawbacks, and therefore, call not only for some defence if they are to be retained, but also for certain improvements if current criticisms are to be met They are, however, equally convinced that in recent years criticism has been pressed too far. At bottom, as Dr. Wiseman observes, the objections to selection, and particularly to selections by means of intelligence tests, are " social and political in origin rather than educational and psychological." But the familiar arguments have been strongly supported by sociological writers, like Dr. Floud and Dr Halsey. In their book on *Social Class and Educational Opportunity* they assure us that " measured intelligence is well known to be largely an acquired characteristic," and therefore unfair to children of the " unprivileged classes " But this, says Dr. Wiseman (quoting Dr Warburton), is rather like declaring that " the English Channel is well known to be wider than the Atlantic Ocean " No one, he adds, who has really studied recent research on the subject " can possibly deny the very substantial genetic component in tested intelligence " ; and, as actual records show, the introduction of such tests has undeniably helped to do fuller justice to those who are socially or culturally handicapped " The intelligence test *must* be used if we are to get anywhere near equality of opportunity ; and we would be more successful in reaching this objective if the test had greater weight."

Professor Oliver defends much the same point of view. From Plato onwards, so he reminds us, philosophers have maintained that educational systems have a dual function—to train and to select ; and there is the inevitable temptation to over-emphasize one function to the neglect or impairment of the other. Current discussion, too, has tended to forget that efficient selection implies and facilitates social mobility. Accordingly, in the interest of the community as well as of the individual, it is to be hoped (he says) that ill-informed attacks on examinations and selection tests will be stoutly resisted. The Crowther Report " tries to escape the necessity of selection by reviving the supposed distinction between selection and qualification." The G.C.E has come to serve both purposes; but the Report proposes to revert to a separate examination for each. Professor Oliver holds that the distinction is unrealistic: if the number qualified exceeds the number of places, we must select; and we ought, so far as we can, to select the best. Hence, he argues, the Crowther Committee " would have done better to address itself to the improvement of the G.C.E —the only public examination at this level which has achieved national currency."

How such improvements might be achieved is admirably outlined by Dr Wiseman Hitherto, he believes, " educational psychologists have tended to ignore the wider educational implications of their work, while the teacher-examiners of the traditional boards do just the reverse " : they overlook the need and the possibility of increasing the sheer efficiency of their techniques In the United States, " comparable boards have research sections engaged full-time on such problems " ; and he rightly urges that there is pressing need for similar studies over here He sums up the whole book by saying that "good examinations are both useful and desirable—without them education would be much less effective than it is : let us, therefore, make sure that the examinations we do use are efficient in the fullest sense."

CYRIL BURT.

WILSON, HARRIET (1962). *Delinquency and Child Neglect.* George Allen and Unwin, pp. 195, 25s.

There have been numerous studies of the relationship between family background and juvenile delinquency, but few of these have been studies in depth of the ' problem family ' in which the children suffer some measure of neglect. Dr. Harriet Wilson's thoughtful contribution to our knowledge in this field is, therefore, to be welcomed and merits careful study. Basing her book on an enquiry into a group of problem families in an industrial city, she sets out to define what is meant by the ' problem family ' and to discover to what extent such families are the breeding ground of delinquency.

After discussing recent trends in studies of juvenile delinquency (she surprisingly omits reference to some important work of relevance, for example that by Stott, in this country, and Hewitt and Jenkins in the United States) and making some interesting comparisons between British and American culture-patterns, Dr. Wilson goes on to describe the patterns of living of fifty-two problem families, selected as her sample on the basis of performance-inadequacy, under the headings of solvency, health and education. She finds that the most disabling feature of the problem family is its social isolation, while emphasizing that these families are not, as is commonly assumed, a homogeneous entity. Delinquency is part of the pattern of juvenile behaviour among families exhibiting signs of child neglect, the main factors dominating the children's lives being poverty and weak relationships with their parents and other adults. The chief characteristic of the juvenile delinquent from the problem family is his complete unawareness of the moral significance of his actions.

The book is refreshingly free from sociological jargon, and tries to avoid unwarranted generalisations. Dr. Wilson has kept an open mind throughout her enquiry and is, on the whole, careful about methodological considerations, though the children have been studied less systematically than the parents. For example, a number of conclusions about the level of the children's intelligence are drawn from data which are incomplete and, being partly based on group tests given at the infant school stage, hardly reliable ; and the author's reason for obtaining teachers' behavioural assessments in an unstandardized form seems inadequate. Furthermore, can one justifiably infer the presence or absence of maladjustment from the correlation between I.Q., attendance, and educational attainment (p. 73) ?

A slight re-arrangement of the material might have made the main findings and conclusions stand out more clearly from the mass of detail; and some of the tables are unnecessarily obscure There is, however, likely to be little disagreement with Dr. Wilson's main conclusions and with her practical proposals in the final chapter, where she advocates, as preventive action, measures which would help the father with a low wage and large family (so that he would no longer be financially little better off working than if unemployed) as well as aid those families permanently dependent on assistance from the state; realistic family planning methods; an increase in the number of nursery schools; and the provision by local authorities of specialized supportive help of a personal kind. MAURICE CHAZAN.

YATES, AUBREY J. (1962). *Frustration and Conflict.* London: Methuen, pp. x+236. 28s.

The author states that this book is not intended to be " an original contribution with respect to theory," but rather " a systematic examination of the accumulated knowledge in the areas of frustration and conflict." These topics have been extensively investigated by psychologists, both clinically and, in more recent years, experimentally. Dr. Yates is concerned primarily with work of the latter kind, but a most useful bibliography of well over three-hundred items shows how widely he has, nevertheless, cast his net.

The book comprises a brief introduction and five main chapters The first is a very thorough review and critical analysis of the work of Maier and his colleagues on frustration and fixation. The author frequently comments on the neglect of Maier's findings by other psychologists and remarks. " There is no doubt that the lasting value of his work has been obscured by sheer prejudice and *a priori* convictions." He, nevertheless, admits that Maier's results and theoretical interpretations are still difficult to reconcile with other contemporary approaches. The second chapter provides a detailed examination of the frustration-aggression hypothesis of Dollard, Miller, *et al.* and of the very considerable number of experiments which this has provoked The author's verdict is that work in this field has proved " extremely profitable." Next comes an account of studies on frustration and regression It includes an exhaustive critique of the famous experiment by Barker, Dembo and Lewin. Dr. Yates notes how extraordinarily few attempts have been made to repeat their findings, a need which is surely pressing, since the original enquiry was conducted without a control group and the statistical treatment of data was often unsatisfactory. The fourth chapter is on conflict. Short sections are devoted to the work of Luria, Lewin and Cattell, but the main analysis is concentrated on the extensive work of Miller and his colleagues. Some most ingenious experiments are discussed in this section, including a particularly brilliant *tour de force*, by Murray and Berkun.

The final chapter is entitled " Frustration, Conflict, Learning Theory and Personality." Here Dr. Yates tries to draw the strands together, and at once the influence of Hullian learning theory is apparent. Recent efforts to incorporate ' frustration ' into learning theory as an intervening variable are described, as are many experiments relating to these attempts. Eysenck's theory of personality, itself of course, very considerably influenced by Hullian theory, is also favoured, and hypotheses are put forward linking frustration and conflict with neuroticism and introversion-extraversion.

As a whole, this book is to be highly recommended. Whatever one may feel about the author's theoretical leanings, serious attempts to test rigorously theories about frustration and conflict are of very recent origin and should be encouraged. Dr. Yates' foremost contribution has been to provide a careful, objective and wide-ranging digest of this interesting, and often exciting, current research For the educationalist, a useful feature is his detailed consideration of the effects of punishment, which are shown to vary greatly under different conditions. More generally, this book provides an essential basic reference both for those contemplating serious research into frustration and conflict and for all advanced students of these topics

G. W. PILKINGTON.

Some Important Books

THE DEVELOPMENT OF THE INFANT AND YOUNG CHILD
By Professor R. S. Illingworth, M.D., F.R.C.P., D.P.H., D.C.H.
324 pages, 95 illustrations. 27s. 6d.

THE NATURAL DEVELOPMENT OF THE CHILD
Fourth Edition. By Agatha H. Bowley, Ph.D., F.B.Psy.S.
224 pages, 84 illustrations. 10s. 6d.

THE PSYCHOLOGICAL CARE OF THE CHILD IN HOSPITAL
By Agatha H. Bowley.
48 pages, 8 illustrations. 4s. 6d.

CEREBRAL PALSY IN CHILDHOOD AND ADOLESCENCE
A Medical, Psychological and Social Study.
Edited by Professor J. L. Henderson, M.D., F.R.C.P.(Ed.). 411 pages. 35s.

E. & S. LIVINGSTONE, LTD.
TEVIOT PLACE, EDINBURGH

THE
BIRMINGHAM
PRINTERS LTD
PRINTERS & BOOKBINDERS

High Class
Technical
Publications

42-44 HILL STREET
BIRMINGHAM 5
TELEPHONE: MIDLAND 0196-7

Educating Exceptional Children

SAMUEL A. KIRK

Director, Institute for Research on Exceptional Children, University of Illinois.

Various kinds and degrees of exceptionality are considered—mental, physical, and those caused by emotional disturbances. The conclusion arrived at by this leading authority will be of the utmost importance to educational psychologists, to teachers in special schools, and to students in training.

432 *pages,* 28 *graphs and figures.* 37*s.* 6*d.*

Speech Difficulties in Childhood

A Practical Guide for Parents and Teachers

RONA M. WILLIAMS *Licentiate of the College of Speech Therapists*

" Lucidly, simply, and yet without any loss of accuracy, Miss Williams describes the whole range of speech difficulties in childhood and very persuasively calls on teacher and parent to play their parts in helping children to overcome their disabilities. There are few teachers whose range of professional understanding will not be broadened by reading this book." PROFESSOR M. M. LEWIS in *The Teacher.*

Illustrated. 15*s. net.*

HARRAP *182 High Holborn, London, W.C.1*

THE BRITISH JOURNAL OF MEDICAL PSYCHOLOGY

Edited by Dr. THOMAS FREEMAN

(*Contents of Vol.* XXXVI, *Part* 1, *March,* 1963)

D. W. Winnicott. Regression as Therapy.
John L. Cameron. Patient, Therapist and Administrator.
Harold F. Searles. Discussion of Dr. Cameron's Paper.
David M. Kissen. Personality Characteristics in Males Conducive to Lung Cancer.
Isabelle V. Kendig. The Denial of Illness.
Stephen A. Appelbaum. The Pleasure and Reality Principles in Group Process Teaching.
S. T. Hayward. On Being Ill in a Mental Hospital.
James M. Davie. Observations on Some Defensive Aspects of Delusion Formation.
Helm Stierlin. Treatment Dilemmas with Psychotic and Sociopathic Patients.
M. C. Moss and P. Hunter. Community Methods of Treatment with Chronic Psychotics.
Reviews.

Subscription price 75s. net per volume of 4 parts. Single parts 25s. net.

Orders may be placed with any bookseller, or sent direct to :

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200, Euston Road, London, N.W.1

STUDIES IN INDIVIDUAL DIFFERENCES

Edited by James J. Jenkins and Donald G. Paterson

'The Search for Intelligence' is the sub-title of this book, which consists of 66 research articles on the subject. It shows how research has developed from its beginnings to the present.

63s.

SCHOOL LEAVERS

Edited by Thelma Veness

An enquiry was carried out among several hundred boys and girls to find the answers to such questions as 'Are the young ambitious? What do they want of life today?' Quotations from essays written show their views on such topics as leisure, marriage and retirement.

With 8 half-tone plates **25s.**

METHUEN *& Co. Ltd. · 36 Essex Street · Strand · London W.C.2*

TAVISTOCK PUBLICATIONS

Educating Older People

M. F. CLEUGH

A study of the social and psychological factors that must be taken into account in teaching adults. Examples illustrating aspects of the teaching situation are discussed as a basis for the formulation of some general principles in the education of the older student.

28*s. net.*

A Textbook of Health Education

DENIS PIRRIE and A. J. DALZELL-WARD

With contributions by Winifred Warden and D. Lynton Porter

The first comprehensive introduction to the general principles of health education designed to provide teachers, students, and public health workers with a basic understanding both of methods and media of instruction and of the biological and social factors underlying problems of community health.

Illustrated. 35*s. net.*

Educational Strategy for Developing Societies

A Study of Educational and Social Factors in Relation to Economic Growth

ADAM CURLE

An approach to the problems of development of the emergent countries based on the hypothesis that the roots of underdevelopment are human as much as economic, social as much as material, and that advance will most effectively result from a well-planned strategy of educational expansion.

25*s. net.*

The Experience of Reality in Childhood Schizophrenia

AUSTIN M. DES LAURIERS

A study based on clinical research concerned with the development of a systematic approach to the psychotherapy of schizophrenia. This work is of significance to two groups—child psychologists and psychiatrists on the one hand, and psycho-analysts and those interested in psycho-analytic theory on the other.

35*s. net.*

11 New Fetter Lane, London, E.C.4.

OTHER PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology.
The British Journal of Psychology (Monograph Supplements).
The British Journal of Medical Psychology.
The British Journal of Statistical Psychology.
The British Journal of Social and Clinical Psychology.

The British Journal of Psychology and *The British Journal of Medical Psychology* appear in quarterly parts in volumes of about 350 pages, subscription price of each, payable in advance, 60s. per volume (post free). Subscriptions may be sent to any Bookseller or to the Cambridge University Press, Bentley House, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $10.00. Enquiries should be addressed to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22.

The British Journal of Statistical Psychology is issued twice a year, in May and November, in volumes of about 184 pages. The subscription rate payable in advance is 30s. 6d. per volume (post free). The subscription price in the U.S.A. is $4.50 (post free). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, E.C.4., or to any Bookseller.

The British Journal of Social and Clinical Psychology is issued three times a year in February, June and November, in volumes of approximately 240 pages. The subscription price of each, payable in advance, is 50s. (post free). Subscriptions may be sent to any Bookseller, or to the Cambridge University Press, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $8.50. Enquiries should be sent to the Cambridge University Press American Branch, 32, East 57th Street, New York 22. Separate parts will be sold at 20s.: $3.50 in the United States.

A series of *Monograph Supplements* is issued in connection with the *British Journal of Psychology*, but is not included in the subscription for that journal. Orders may be sent to any Bookseller or to the Cambridge University Press, 200, Euston Road, London, N.W.1., from whom details of these publications may be obtained.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to:
Dr. BORIS SEMEONOFF, Department of Psychology, The University, Edinburgh.

For the *British Journal of Medical Psychology*, should be sent to Dr. THOMAS FREEMAN, Lansdowne Clinic, 4, Royal Crescent, Kelvingrove Glasgow, C.3.

For the *British Journal of Educational Psychology* should be sent to Mr. L. B. BIRCH, Institute of Education, The University, Sheffield, 10.

For the *British Journal of Statistical Psychology*, should be sent to the Assistant Editor, Sir CYRIL BURT, 9, Elsworthy Road, Hampstead, London, N.W.3.

For the *British Journal of Social and Clinical Psychology* should be sent to Mr. MICHAEL ARGYLE (Social Psychology Editor), Institute of Experimental Psychology, 1, South Parks Road, Oxford;

or to:

Dr. JACK TIZARD (Clinical Psychology Editor), Maudsley Hospital, Denmark Hill London S.E.5.

For the *Monograph Supplements*, to Professor L. S. HEARNSHAW, Department of Psychology, The University, 7, Abercromby Square, Liverpool, 7, whose decision regarding the publication of all *Monograph Supplements* will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1st September of the preceding year.

OFF-PRINTS

Contributors receive 50 free copies of their article and may buy more in units of 50.

Members of the British Psychological Society receive all the Society's Journals at reduced rates.

Information on membership and further activities may be obtained from the Secretary, Tavistock House South, Tavistock Square, London, W.C.1.

PUBLISHED BY THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May and distributed free to members of the Association. To non-members it costs, 5s. per issue, post free (subscription 14s. per annum).

Members of the A.T.C.D.E. receive the *British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151, Gower Street, London, W.C.1.

CONTENTS

PAGE

K. Lovell, C. Byrne and B. Richardson: A Further Study of the Educational Progress of Children who had received Remedial Education . 3

M. Woodward and Diana J. Stern: Developmental Patterns of Severely Sub-Normal Children 10

C. Anne Moreton and H. J. Butcher: Are Rural Children Handicapped by the Use of Speeded Tests in Selection Procedures? 22

T. G. Crookes and Margaret C. L. Greene: Some Characteristics of Children with Two Types of Speech Disorder 31

H. J. Hallworth: An Analysis of C. W. Valentine's Reasoning Test for Higher Levels of Intelligence 41

G. Jahoda: The Development of Children's Ideas About Country and Nationality: I.—The Conceptual Framework . 47

G. Harry McLaughlin: Psycho-Logic: A Possible Alternative to Piaget's Formulation 61

F. W. Warburton, H. J. Butcher and G. M. Forrest: Predicting Student Performance in a University Department of Education 68

Critical Note:

D. Russell Davis and Asher Cashdan: Specific Dyslexia 80

M. D. Vernon: A Reply to: Specific Dyslexia, by R. Russell Davis and Asher Cashdan 83

R. R. Dale: A Brief Comment on: The Effect of Socio-Economic Status on the Scholastic Achievement of Sixth-Grade Children, by Robert L. Curry 84

Research Note:

L. R. Bowyer, A. Marshall and K. Weddell: The Relative Personality Adjustment of Severely Deaf and Partially Deaf Children 85

Summaries of Researches Reported in Degree Theses 88

Book Reviews 95

22 APR 1963

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

JUNE 1963

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

VOL. XXXIII, PART 2

Price One Pound

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
L. B. BIRCH

Associate Editors:
P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
SIR FRED J. SCHONELL
C. W. VALENTINE
S. WISEMAN

Vol XXXIII, Part 2

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education, once a term, in February, June and November. The subscription is £2 a year post free payable in advance, or £1 each number.

Subscriptions and Orders should be sent to "The Manager, B.J.E.P. Department, Methuen and Co., Ltd., 36, Essex Street, Strand, London, W.C.2." Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to L. B. Birch, The Institute of Education, The University, Sheffield, 10. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to L. B. Birch, at the above address. *Advertisements* and *Subscriptions* for this *Journal* should be sent to Manager, B.J.E.P. Department, Methuen and Co., 36, Essex Street, Strand, London, W.C.2.

Published by
METHUEN AND CO., LTD., 36, ESSEX STREET, STRAND, LONDON, W.C.2.,
for the British Journal of Educational Psychology, Ltd.

Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

VOLUME XXXIII .JUNE, 1963 PART 2

THE REDUCTION OF SOCIOECONOMIC BIAS IN INTELLIGENCE TESTING

BY R. S. MACARTHUR AND W. B. ELLEY
(University of Alberta)

SUMMARY. This study was designed to investigate the extent of socioeconomic bias in selected intelligence tests, with a view to identifying and constructing instruments which might better estimate the intellectual potential of pupils from deprived cultural and educational backgrounds. For such pupils the conventional intelligence test is unsuitable as a predictor of later educational progress if adaptive treatment reduces the effects of their educational handicaps.

Nine promising culture-reduced tests and subtests were administered to a representative Edmonton sample of 271 twelve and thirteen-year-old children, and the results analysed in relation to measures of socio-economic status, verbal intelligence, achievement and school marks. City-wide testing surveys conducted during the previous four years provided data on the same children, making longitudinal comparisons possible. The results supported the hypothesis that it is possible to measure a broad component of intellectual ability with significantly less cultural bias than is found in the conventional intelligence test. Raven's Progressive Matrices proved the most useful test in the battery since it showed high g loading, consistent and minimal relation with socioeconomic status, no evidence of cultural bias by items, and moderate correlations with school marks. The Cattell test of g and the Lorge-Thorndike Figure Analogies subtest showed similar promise, since they met most of the criteria of the study in shorter testing time.

I.—INTRODUCTION

INTELLIGENCE tests are frequently evaluated by mechanical investigation of their predictive validity in relation to academic criteria. Usually underlying such an empirical approach is the assumption that only one or two pre-arranged and rather standard treatments will occur between test and criterion judgments.

Cronbach and Gleser (1957) have drawn attention to a distinction between fixed-treatment and adaptive-treatment situations in decision-making based on the results of mental tests. In the former case, treatment is specified *a priori*, and maximum utility is related to predictive validity coefficients ; individuals are selected or placed simply in accordance with their promise in relation to a predetermined criterion with predetermined treatments, and little regard for the qualities underlying the test is necessary. In educational prediction, the intervening treatments are usually limited by a relatively fixed curriculum and relatively standard teaching techniques.

In adaptive teaching decisions, however, atypical individuals may be assigned quite different treatments according to their present status and indicated potential. In an educational setting, immigrant children may be given supplementary English coaching, children from deprived backgrounds may benefit from remedial reading, and sensory-handicapped children may be offered special aids. Educational philosophies which emphasize the development of potential to its maximum require serious consideration of adaptive rather than fixed treatments, and to aid in such adaptive treatment, the use of mental tests which can provide estimates of potential ability, divorced as much as possible from inequalities in cultural background. The conventional verbal test of intelligence, designed as it is for subjects raised within a specific cultural range, has limited suitability for identifying potentially gifted children from deprived or foreign backgrounds, or with restricted formal education. The writers of this article suggest that various non-verbal tests for use in adaptive treatment settings have received insufficient research attention.

It is further suggested that the choosing of a test to help identify potential for attainment of distant and general criteria, assuming the possibility of intervening treatments appropriately adapted to realize that potential, can be facilitated through the ' construct validity ' approach advocated by Cronbach and Meehl (1955). The evaluation of tests of educability should place more emphasis than is customary on theoretical analysis of their relation to the concept of intelligence as developed by experience, with proportionately less emphasis on their ' cold-blooded ' predictive validity under limited fixed treatments. If potentially suitable tests can be examined in the light of hypotheses generated from an adequate theory of intelligence, some progress may be anticipated in the development of culture-reduced tests suitable for classifying and selecting children from diverse cultural backgrounds.

It was the purpose of the present investigation to identify tests or subtests which, for a sample of 12 and 13-year-old pupils from a variety of socioeconomic backgrounds in Edmonton, Canada, could on the one hand, be regarded as valid measures of intellectual ability, while on the other hand, could demonstrate a minimum of socioeconomic bias. Parallel studies with children from different ethnic and language backgrounds are now in progress (MacArthur, 1962).

II.—PREVIOUS RESEARCH.

Binet himself realized that his measuring scales were sensitive to ' scholastic and extra-scholastic acquisitions ' as well as ' intelligence pure and simple ' (1916). Further analyses by Stern (1914) and Burt (1921) showed significant social class differences, particularly in the more verbal items of the Binet scales, while Gordon (1923), Wheeler (1942), Baldwin, *et al* (1930) and others have shown that the conventional intelligence test was unsuitable for children without formal education, from isolated communities or from rural districts in general.

Amongst the causal interpretations made of social class differences in measured intelligence major emphasis has usually been placed on hereditary and environmental influences. While not denying these, it is considered that the importance of test bias requires further study. Eells (1951) has produced strong evidence to support the contention that bias in intelligence test items is common, and that it tends to favour upper-middle class children. From his Chicago surveys, he concluded that verbal tests and items show significantly more variation in discriminating between social classes than do non-verbal.

Havighurst, *et al* (1944, 1947) also demonstrated that the extent of differentiation amongst social classes depends on the trait measured, the greatest discrepancies appearing in verbal and numerical tests, the smallest in spatial, mechanical and other non-verbal tests.

These research studies are, unfortunately, open to the criticism made by Lorge (1952) that analysis in terms of social levels reveals nothing about the ability of tests to differentiate between intellectual levels. Both dimensions are important in test evaluation. Moreover, Eells used only one non-verbal test whose reliability was suspect (Tyler, 1953), while Havighurst's studies were based on relatively small samples.

'Culture-fair' tests such as those of Cattell (1940, 1958) and the Davis-Eells Games (1953) have not adequately fulfilled their designers' intentions according to studies described by Anastasi and Cordova (1953), Drake (1955), Smith (1956) Stilwell (1958) and Love and Beach (1957), although the evidence has sometimes been more favourable (Marquant and Bailey 1955, Cattell 1958, Fowler 1955). Further evaluation of such tests using a construct validity approach in a variety of settings seems necessary in order to assess the usefulness of such measures.

Since the present investigators are seeking to identify tests suitable for a variety of cultural and educational backgrounds, it was considered that Cattell's approach to the problem of reducing cultural differentials by holding them to a minimum would be more suitable for this study than the Davis-Eells approach of holding them constant. On such an assumption, this investigation was designed to study tests similar to Cattell's in an attempt to identify test items which tap fundamental processes of intellectual ability while holding environmental differences to a minimum.

No illusions are held about the possibility of eliminating cultural influences from a pencil-and-paper test. Every culture stimulates the development of certain abilities at the expense of others, and the idea of an intelligence test being culture-free in the sense that it measures the products of heredity quite apart from the impact of cultural experience cannot be taken seriously. The present study aims rather to identify and evaluate tests which reduce the dependence on acquired knowledge in order to make possible the improvement of mental testing for children whose backgrounds do not fall within the required range assumed by the conventional verbal test.

III.—THE CONCEPT OF INTELLIGENCE.

Over the past several decades the concept of intelligence as general innate ability has been subjected to a barrage of criticisms. Factor analysts such as Thurstone (1938) and Guilford (1956) have disparaged the idea of a general factor; environmentalists such as Halsey (1959) and the Iowa investigators Stoddard and Wellman, 1940) have questioned the notion that intelligence is predetermined or fixed; developmental research pioneered by Piaget (1947) suggests that intelligence is intimately bound up with the continuous interaction between organism and environment, while the possibility of critical periods and irreversible influences in mental development is now seriously held (Hebb, 1958).

One beneficial effect of these criticisms has been to draw attention to the distinctions drawn by Hebb (1949) and Vernon (1955) amongst Intelligences A, B and C. Nevertheless, behavioral definitions of intelligence are as controversial as ever, and the operational approach of defining intelligence in statistical terms appears the more profitable course of action. Such a position was adopted in forming the conceptual framework of this study.

The hierarchical conception of intelligence described by Burt (1949) and Vernon (1950) is regarded by the present authors as the most useful model for assessing intelligence in diverse cultural settings. The general factor is a crucial construct for assessing the extent to which culture-reduced tests are sampling the common denominator of abilities required for success in a variety of mental tests. It looms so large in psychometric surveys of heterogeneous children that it is difficult to ignore, while the advantage of distributing its variance over a number of lower-level factors has yet to be demonstrated even with culturally homogeneous children. In predicting to distant and general criteria, when adaptive treatments are possible, 'pure' factor tests, are seldom superior to general intelligence tests, while the amount of time and money consumed on the former is an obvious administrative disadvantage.

The question of the innateness of the general factor can be circumvented by insisting only on its comparative stability in relation to lower-order factors. Evidence to support such stability comes from various sources. Woodrow (1938) demonstrated that the general factor was relatively unaffected by differential practice in a thirty-three variable analysis, while group factors changed considerably. Hofstaetter's (1954) factor analysis of the longitudinal mental development of children in the Berkeley Growth Study revealed a stable general factor increasing in size between the ages of 4 and 8, and remaining steady throughout pre-adolescence. Tests with high g loadings and negligible group factors have shown less susceptibility to cultural change (Cattell, 1958) and other environmental influences (Nisbet 1953, Burnett 1958).

Theoretically, the stable g factor is compatible with the hypotheses of Hebb (1949) and Ferguson (1956) which suggest that abilities are developed primarily by positive transfer from the old learning to the new. Early superiority in processing and integrating skills should lead to increasingly broad and cumulative superiority of intellectual development in a variety of fields. Vernon (1960), too, sees support for the general construct in the developmental approach for " those who, early in life acquire a larger stock of perceptual schemata and verbal habits are better able to build up more complex and more flexible schemata necessary for conceptual thinking " (p. 41). So the general all-inclusive logical structures accumulate and proliferate, producing a broad general factor which, although difficult to define behaviourally, continues to account for most of the statistical variance in any battery of ability tests.

In a comprehensive review of theories and research on intelligence, Hunt (1961) draws attention to the compatibility of apparently diverse interpretations of intelligence, which seem to be best explicable in hierarchical terms.

" A hierarchical factor structure, with Spearman's g at the apex, as the English factor analysts have tended to conceive of intelligence is precisely what would be expected from considering abilities as the accumulative effects of the transfer of training or learning in varied situations . . . Piaget's conception of accommodation corresponds to the change in function that occurs with training or learning, and Piaget's conception of assimilation corresponds to the notion of transfer of training " (Hunt, 1961, p. 359).

This general factor at the top of the hierarchical model, is essential to the present study. To qualify as a measure of an all-embracing intellectual ability, while relatively free from the vagaries of environmental circumstances, a test must show a substantial loading on the general factor at the top of the hierarchy. Culture reduction is of little value if the test does not sample extensively that common denominator or ' transfer basis ' of the fluid set of abilities we understand by intelligence.

IV.—EXPERIMENTAL PROCEDURE.

Nine tests and subtests were selected as promising culture-reduced measures of general intelligence. In May, 1960, these were given to 271 pupils constituting a stratified sample from the Grade VII population in the public schools of Edmonton, a rapidly expanding central Alberta community of nearly 300,000 people. The variables of provincial examination averages, socioeconomic status, geographic location, school type and residential mobility were considered in selecting the sample, and a comparison with a random selection from the same population in terms of age, sex, occupational status and Laycock I.Q., showed no significant differences.

The potentially useful culture-reduced tests selected were :
(1) Raven Standard Progressive Matrices (1956).
(2) Cattell Test of g " Culture Free." Scale 2. (1958).
(3) Lorge-Thorndike Non-Verbal Intelligence Test. Level 4 (1957).
(4) Lorge-Thorndike Figure Classification Subtest.
(5) Lorge-Thorndike Number Series Subtest.
(6) Lorge-Thorndike Figure Analogies Subtest.
(7) Holzinger-Crowder Uni-Factor Test (1952) Figure Changes.
(8) Holzinger-Crowder Uni-Factor Test (1952) Mixed Series.
(9) California Test of Mental Maturity (CTMM) Non-Language, Short Form, Elementary (1957).

Other variables for which data were gathered were :
(10) California Test of Mental Maturity, Language (1957).
(11) Holzinger-Crowder Uni-Factor Spatial Test (Boots) (1952).
(12) Occupational Index Score.
(13) Home Index Score.
(14) Year-end marks on teacher-made tests of attainment.

In May of the previous year, 1959, the same pupils had been administered the California Achievement Battery, Junior High (1957), whose subtests, listed below, were used to help identify factors in the factor analytic studies.
(15) California Reading Vocabulary.
(16) California Reading Comprehension.
(17) California Reading Total.
(18) California Arithmetic Reasoning.
(19) California Arithmetic Fundamentols.
(20) California Arithmetic Total.
(21) California Language.
(22) California Spelling.
(23) California English Total.

Further data used for comparative purposes were :
(24) Laycock Mental Ability Test (1935)—raw scores obtained two years earlier.
(25) Raven Coloured Progressive Matrices (1947)—administered in May, 1956.
(26) CTMM Primary (1953)—administered in May, 1956.

The CTMM Elementary (1957) was also analysed by subtests.
(27) CTMM Spatial.
(28) CTMM Logical.
(29) CTMM Number.
(30) CTMM Verbal.

The occupational index was obtained by assigning each pupil's statement (checked in a local directory) a standard score from Blishen's (1958) Canadian Socio-Economic Scale, based on Dominion census data, and evenly weighted according to income and years of training required. The Home Index Scale consisted of twenty items, adapted for local conditions from a scale devised by Gough (1949). The purpose of these measures was not to isolate the crucial variables operating to accelerate mental maturity, but rather to indicate when a number of these stimulating correlates was present and operating directly or indirectly, and when their importance was reduced.

On the basis of previous research, and the theoretical position adopted, the following predictions were made :

(*a*) The nine culture-reduced tests will show g loadings as high as those of the conventional verbal tests of intelligence.

(*b*) The nine culture-reduced tests will show negligible loadings on those group factors which appear to depend heavily on cultural acquisitions.
(*c*) Socioeconomic status of father will show greater correlations with scores on the conventional verbal tests than it does with scores on the culture-reduced tests.
(*d*) Over a period of four years (1956 to 1960) children from upper socioeconomic levels will improve their status on the CTMM more than on the culture-reduced Progressive Matrices.
(*e*) An item analysis of the most promising culture-reduced tests will show that upper and lower thirds, by socioeconomic status, will succeed in similar ratios on each item. On the CTMM Verbal Test, discrimination indices will show greater variation from item to item.

Additional evaluative criteria of relationship with school achievement, reliability and suitability for classroom use were also examined.

V.—ANALYSIS OF RESULTS.

Three socioeconomic variables were used in the analyses—occupational T-score, home index raw score, and a combination of the two measured with equal weights, yielding what was termed a socioeconomic index. Preliminary analysis on the two component sub-scores showed an intercorrelation of 0·61 which is close to Gough's corresponding coefficient of 0·65 (1949). The pattern of correlations with other variables was almost identical for the two measures (rho=0·87) suggesting that they sample highly similar correlates of mental ability. A split-half reliability coefficient of 0·77 was obtained from the Home Index questionnaire.

(*a*) and (*b*) The two hypotheses concerning the abilities sampled by each test were investigated by performing two factor analyses using Hotelling's Principal Components Method. The first analysis was performed on twenty-one variables (Tables 1 and 2) and included many subtests. The second concentrated on the eleven total test scores (Tables 3 and 4).

From the first matrix six principal components were extracted* but the last was discarded because it failed to reach significance level, according to the tests of McNemar, Burt, and Guilford and Lacey. The five significant factors were then rotated using criteria of positive manifold, maximization of the general factor and psychological meaningfulness. After four rotations a compelling psychological factor pattern emerged (Table 2). The factors were interpreted as general ability (g), a verbal educational factor (v : ed), socioeconomic status (se), numerical ability (n) and, possibly, speed (sp). The fact that no spatial factor appeared in the analysis may be due to the low ceiling of the CTMM Spatial Subtest, or the strong speed element in the Holzinger-Crowder Test. The eleven total test scores were analysed similarly yielding three clear factors labelled g, v : ed and n. A comparison of the factor patterns of the tests common to both analyses showed a mean loading discrepancy of only ·06, which suggests a high degree of similarity in the common dimensions obtained. The rotated factor solutions indicate that the tests most highly loaded with the general factor were Cattell's Test of g, the Progressive Matrices, and the Lorge-Thorndike Figure Analogies Subtest. The remaining six culture-reduced tests showed less promising g saturations, or significant loadings on verbal or numerical factors.

* Communalities were estimated by calculating the squared multiple correlation of each variable with the n-1 remaining variables in the battery, as suggested by Wrigley (1959) and Kaiser (1960). Iteration was performed after replacing the communality with a value increased (or decreased) by one-and-a-half times the discrepancy between obtained and estimated values, until convergence occurred.

TABLE 1

Correlation Matrix for Twenty-one Variables.
(Decimal Points Omitted.)

Test No.	1	2	4	5	6	7	8	11	12	13	15	16	18	19	21	22	24	27	28	29	30
1	(542)*																				
2	637	(629)																			
4	439	441	(358)																		
5	457	422	387	(494)																	
6	566	555	485	403	(556)																
7	494	457	347	388	379	(442)															
8	358	387	338	532	396	451	(555)														
11	417	393	237	247	262	356	277	(328)													
12	165	189	113	181	240	143	239	081	(552)												
13	237	244	152	176	233	237	302	168	610	(582)											
15	301	254	225	334	333	334	427	176	265	328	(742)										
16	466	387	280	405	405	427	461	161	250	303	681	(645)									
18	349	360	308	368	365	316	377	215	177	234	467	465	(371)								
19	396	303	335	435	362	358	470	130	118	162	447	586	459	(621)							
21	404	341	296	363	340	405	500	224	179	268	656	653	427	532	(647)						
22	172	068	105	143	207	114	193	−037	132	154	492	438	330	425	450	(445)					
24	507	498	435	438	526	428	451	268	309	324	643	662	489	479	621	441	(740)				
27	502	465	373	369	471	284	342	310	059	060	145	213	287	233	250	026	308	(437)			
28	527	493	362	383	455	372	373	284	286	239	355	459	313	317	372	184	550	382	(483)		
29	538	466	318	510	448	425	471	228	276	282	364	405	418	510	402	276	554	349	517	(571)	
30	337	352	291	340	374	353	425	136	286	383	717	503	438	421	580	449	667	187	469	429	(700)

* Final estimate of communality.

TABLE 2

ROTATED FACTOR SOLUTION AND COMMUNALITIES FOR TWENTY-ONE VARIABLES.
(Loadings below ·20 omitted)

Test No.	Test Name	I_3	II_2	III_1	IV_1	V_1	L^2
1	Prog Matrices	708	—	—	—	—	542
2	Cattell g	754	—	—	—	222	626
3	L-Th. Fig. Class	582	—	—	—	—	355
5	L-Th. No. Series	547	—	—	395	—	486
6	L-Th. Fig. Anal.	738	—	—	—	—	553
7	Holz -Crowd. Fig Ch. ..	498	—	—	211	369	440
8	Holz.-Crowd. Series	457	214	—	453	303	554
11	Holz.-Crowd Spatial	397	—	—	—	396	326
12	Occ. Status	251	—	683	—	—	583
13	Home Index	247	205	669	—	—	602
15	Reading Vocab.	335	743	—	—	264	747
16	Reading Comp.	501	618	—	—	—	681
18	Arith. Reasoning	460	341	—	—	—	369
19	Arith. Fund.	450	436	—	442	—	629
21	Language	417	589	—	—	299	641
22	Spelling	201	618	—	—	—	443
24	Laycock	676	509	—	—	—	735
27	CTMM Spatial	612	—	—	—	—	428
28	CTMM Logical	660	—	—	—	—	476
29	CTMM Number	639	201	—	331	—	563
30	CTMM Verbal	458	658	—	—	—	694
% of Common Variance		51·18	25·23	9·11	7·80	6·68	100 00
% of Total Variance		27·94	13·78	4·97	4·26	3·65	54·60

TABLE 3

CORRELATION MATRIX FOR ELEVEN TOTAL TEST SCORES.
(Decimal Points Omittted.)

Test No.	1	2	3	7	8	9	10	17	20	23	24
1	(620)*										
2	637	(645)									
3	629	614	(666)								
7	494	457	470	(555)							
8	358	387	531	451	(555)						
9	577	505	605	384	460	(515)					
10	473	473	549	434	492	443	(727)				
17	425	356	467	419	486	351	772	(800)			
20	435	369	532	393	504	425	562	629	(574)		
23	381	297	408	363	466	353	619	733	598	(660)	
24	507	498	601	428	451	431	732	713	550	642	(741)

* Final estimate of the communality.

TABLE 4

ROTATED FACTOR SOLUTION AND COMMUNALITIES FOR ELEVEN TOTAL TEST SCORES. (Loadings below ·20 omitted).

Test No.	Test Name	I_2	II_1	III_1	L^2
1	Prog. Matrices	780	—	—	634
2	Cattell g	792	—	—	639
3	Lorge-Thorndike	740	—	325	678
7	Holz.-Crowd. Fig Ch. ...	521	—	256	374
8	Holz.-Crowd. Series	399	338	531	556
9	CTMM Non-Lang.	624	—	349	516
10	CTMM Language	580	616	—	720
17	Calif. Reading	444	784	—	819
20	Calif. Arith.	430	509	343	562
23	Calif. Eng.	361	702	—	654
24	Laycock	634	573	—	730
	% of Common Variance	55·97	32 83	11·20	100·00
	% of Total Variance	35·02	20·55	7·01	62·57

(*c*) To investigate the third hypothesis, the combined index of socioeconomic status was correlated with each test, and the difference between coefficients tested for significance. Table 5 shows the correlations of socioeconomic status with the nine culture-reduced tests in the left column and four verbal tests across the top. Significance levels of difference between coefficients are found in the body of the table.

Only three of the observed differences failed to reach the ·05 significant level, a finding which supports the hypothesis that the culture-reduced tests selected show less differentiation between socioeconomic levels than do the conventional verbal tests. The three tests found to have highest g loadings showed satisfactory linearity of regression and so can be considered to meet the third criterion of culture-reduction quite adequately.

TABLE 5

SIGNIFICANCE LEVELS OF DIFFERENCES IN CORRELATION COEFFICIENTS BETWEEN SOCIO-ECONOMIC STATUS AND SELECTED TESTS.

Test	Test	CTMM Total	Laycock Total	CTMM Lang.	CTMM Verbal
	r with S E.S.	·384	·354	·414	·382
Progressive Matrices....	·228	·01	·02	·01	·01
Cattell Test of g	·244	·01	·05	·01	·02
Lorge-Th. Fig. Class ..	·152	·01	·01	·01	·01
Lorge-Th. No. Series ..	·192	·01	·01	·01	·01
Lorge-Th. Fig. Anal. ..	·262	·01	·05	·01	·05
Lorge-Th. Total........	·269	·01	·05	·01	·05
CTMM Non-Lang.	·184	·01	·01	·01	·01
Holz -Crowd. Series ..	·310	N.S.	N.S	·05	N.S.
Holz.-Crowd. Fig. Ch. ..	·220	·01	·02	·01	·01

(*d*) The Progressive Matrices and CTMM tests were used to test the fourth hypothesis of change in relationship with socioeconomic status over four years. Since appropriate forms of the Matrices and CTMM had been administered to the sample in both Grades III and VII, it was possible to predict an increase in correlation with socioeconomic status for the more culture-loaded but not the culture-reduced test. The fact that the correlations between grade levels were similar for each test (·545 and ·471, respectively) made the assumption of similar degree of overlap in the functions measured on both occasions appear reasonable, while the same coefficients were considered low enough to allow those socioeconomic variables associated with test scores to operate.

Correlation coefficients with socioeconomic status were calculated for normalized T-scores on all four tests, and differences tested for significance.* Results showed an insignificant increase of 0·015 in the case of the Matrices, whereas the corresponding CTMM Total correlation increased from 0·215 in Grade III to 0·323, a difference which is significant at the ·05 level. As predicted, the abilities sampled by the Matrices test appear less susceptible to the cumulative effects of those socioeconomic variables which influence the mental development of children between Grades III and VII.

(*e*) To investigate the fifth hypothesis, that a culture-reduced test would show less variation than a culture-loaded test in terms of item discrimination indices across socio-economic levels, the Progressive Matrices and CTMM Verbal Test were selected as representative of each category. The percentage passing each item in high and low socio-economic thirds was calculated for each test, distributions of discrimination indices drawn up for each test, and their variances compared. Only items with difficulty indices between 0·10 and 0·90 were included in the analysis.

The variances of the two distributions for the Matrices and CTMM Verbal Tests were calculated as 16·32 and 35·28, respectively, which, when compared, yielded a significant F ratio of 2·16 with 28 and 35 d.f. As hypothesized, items in the verbal test show greater variation in their discrimination between some socioeconomic levels than those of the non-verbal test. Although these discrepancies may not mar the predictive validity of the test assuming fixed treatments, they suggest that the verbal test is less suitable as a measure of intellectual potential. The Matrices discrimination indices were found to be similar throughout a wide range of difficulty.

The criterion of relationship with school marks and standardized tests was then investigated. End-of-year examination averages over all school subjects for each of the thirteen classes in the study were correlated with each of the intelligence tests and subtests. Results from the California Achievement tests administered in the previous year were also correlated with the major variables, using the whole sample of 271 pupils. The median correlations for the thirteen school marks and the mean correlations for the three California Achievement subtests are included in Table 6. Although less efficient than the culturally weighted conventional tests in this ' cold-blooded ' predictive situation, the correlations of the better culture-reduced tests are substantial enough to indicate promise for adaptive treatments.

Typical reliability coefficients reported from previous research are included in Table 6. A check was made on the Progressive Matrices in view of its promise in terms of the other criteria. A split-half reliability co-efficient of 0·91 and a Kuder-Richardson coefficient of 0·83 indicated a degree of reliability very close to the average value reported in other studies at this age level. (Burke, 1958).

---

* The sample was reduced from 271 to 243 to test this hypothesis due to incomplete test data from Grade III.

All tests were considered convenient for classroom use, with the possible exception of the Holzinger-Crowder subtests, whose administration instructions were found to be tedious and repetitive, and time limits somewhat unrealistic. The Matrices test was judged the most popular by pupils and its ease of administration could certainly be regarded as an asset.

VI.—DISCUSSION AND CONCLUSIONS.

The main findings of the study have been drawn together in Table 6. In interpreting this table it should be remembered that, assuming adaptive rather than fixed treatment, the criteria of high g loading and low correlation with socioeconomic status should be accorded more weight than the present relationship with school achievement.

Table 6 indicates that the Progressive Matrices best meets the criteria chosen for evaluation of a culture-reduced intelligence test. In addition to the high g loading and low correlation with socio-economic status, this test is a reasonably good predictor of school marks, has a respectable reliability, shows no change in relationship with socioeconomic status over four years, and gives little indication of variation in item discrimination across socioeconomic levels. Moreover, the test has face validity in that the progressive increase in difficulty makes it a tester of what it teaches and it appears to be little affected by the previous experience of the testee.

The shorter Cattell Test of g and Lorge-Thorndike Figure Analogies Subtest rank next. Reorganization of instructions, lengthening, and item analyses might well improve these tests, particularly in relation to reliability and correlation with schoool marks, where they are weakest. The remaining culture-reduced tests showed less promise in this investigation although further study is currently being conducted on their potential values in relation to racial and language variables.

TABLE 6

SUMMARY OF SELECTED INTELLIGENCE TESTS EVALUATED AGAINST SIX CRITERIA.

Test	g loading	v·ed or n	r with S.E S.	Median r Sch Marks	Mean r Acht. Tests	Split-Half Reliability
Prog. Matrices	·78	—	·23	·56	·41	·91
Cattell g.........	·79	—	·24	·34	·35	·85*
Lorge-Th. Fig. Class	·58	—	·15	·31	·31	·71
Lorge-Th. Series .	·55	·40(n)	·19	·45	·41	·84
Lorge-Th. Fig. Anal	·74	—	·26	·42	·39	·87
Lorge-Th Total ..	·75	·33(n)	·27	·43	·47	·93
Holz.-Crowd. Series	·46	·45(n)	·31	·51	·49	—
Holz -Crowd. Fig.Ch	·52	·21(n)	·22	·40	39	—
CTMM Non-Lang.	·62	·35(n)	·18	·38	·38	·91
CTMM Lang. ...	·58	62 (v·ed)	·41	·67	·66	·95
CTMM Total	—	—	·38	·66	·65	·95
Laycock Total ..	·68	·57 (v:ed)	·35	·59	·64	·95

* Re-test reliability coefficient.

For further research, too, one might predict considerable potential for an unspeeded test which used the Matrices approach of testing what it taught, and which supplemented the abilities measured with the most discriminating figure analogy and perceptual series items, or variations thereof.

In summary, it seems reasonable to conclude that as tests of intellectual potential, the conventional verbal intelligence tests examined in this study show discrimination against children from lower socioeconomic backgrounds. It has been shown that, for 12 and 13-year-old pupils :

(1) Culture reduced tests sample the general intellectual ability factor as well as or better than conventional tests.
(2) Most culture-reduced tests show negligible loadings on verbal and numerical factors.
(3) Culture-reduced tests show significantly less relationship with socio-economic status than do conventional tests.
(4) A conventional verbal test (CTMM) showed significant increase in relationship with socio-economic status over four years, whereas the Progressive Matrices showed no change.
(5) Verbal items from the CTMM showed greater variation in item discrimination between social classes than did items from the Progressive Matrices.

Generalizing further, the three most promising tests identified by the study could profitably be put to greater use in the estimation of intellectual ability for adaptive teaching purposes, and particularly in talent identification and classification of pupils from underprivileged and foreign language bckgrounds, for distinguishing between mental defect and reading difficulties due to other causes, for holding intelligence constant in research experiments where verbal ability is irrelevant, and generally, as a supplement to verbal intelligence measures in the routine testing of children from all socioeconomic levels.

VII.—REFERENCES.

ANASTASI, A., and CORDOVA, F. A. (1953). Some effects of bilingualism upon the intelligence test performance of Puerto Rican children in New York City. *J. Educ. Psychol.*, 44, 1-18.

BALDWIN, B. T., FILLMORE, E A., and HADLEY, L. (1930). *Farm Children.* New York : Appleton-Century-Crofts.

BINET, A., and SIMON, T. (1916). *The Development of Intelligence in Children.* (Translated by Kite, E. S.) Baltimore : Williams and Williams.

BLISHEN, B. R (1958). The construction and use of an occupational class scale *Canad. J. Econ. Polit. Sci.*, 24, 519-531.

BURKE, H. R. (1958). Raven's progressive matrices : Review and critical evaluation. *J. Genet. Psychol*, 93, 199-228.

BURNETT, A. (1955). Assessment of intelligence in a restricted environment. Unpublished Ph D. Thesis. McGill University.

BURT, C. (1922). *Mental and Scholastic Tests.* London : King and Son.

BURT, C. (1949). The structure of the mind : a review of the results of factor analysis. *Brit. J. Educ. Pyschol*, 19, 110-111, 174-179.

CATTELL, R. B. (1940). A culture-free intelligence test. *J. Educ Psychol.*, 31, 161-180.

CATTELL, R. B. (1958). *Handbook for the Culture-Free Intelligence Test, Scale* 2. Champaign, Illinois : IPAT.

CRONBACH, L. J., and GLESER, G. C. (1957). *Psychological Tests and Personnel Decisions.* Urbana, Illinois : University of Illinois Press.

CRONBACH, L. J., and MEEHL, P. E. (1955). Construct validity in psychological tests *Psychol. Bull*, 52, 281-302.

DAVIS, A , and EELLS, K. (1953). *David-Eells Test of General Intelligence or Problem Solving Ability.* New York : World Book Co.

DRAKE, R. M. (1953). Review of Cattell Culture-Free Test, in Buros, O K. (Editor). *Fourth Mental Measurements Yearbook.* New Jersey : Gryphon Press.

Eells, K., Davis, A., Havighurst, R. J., Herrick, V. E., and Tyler, R. S. (1951). *Intelligence and Cultural Differences* Chicago : University of Chicago Press.

Elley, W. B. (1961). A comparative analysis of the socioeconomic bias in selected intelligence tests. Unpublished Ph.D. Thesis, University of Alberta.

Ferguson, G. A. (1956). On transfer and the abilities of man. *Canad. J. Psychol.*, 10, 121-131.

Fowler, W. L. (1955). A comparative analysis of pupil performance on conventional and culture-controlled tests. Ph.D. Thesis, Ann Arbor, Michigan.

Gordon, H. (1923). Mental and scholastic tests among retarded children. London : *Bd. Educ , Educ. Pamphlet, No. 44.*

Gough, G. (1949). A short social status inventory. *J. Educ. Psychol.*, 40, 52-56.

Guilford, J. P. (1959). *Personality*. New York : McGraw-Hill.

Halsey, A. H (1959). Class differences in general intelligence : I. *Brit. J. Stat. Psychol.*, 12, 1-4.

Havighurst, R. J., and Breese, F. H. (1947). Relation between ability and social status in a mid-western community : III. *J. Educ. Psychol.*, 38, 241-247.

Havighurst, R. J., and Janke, L. L. (1944). Relation between ability and social status in a mid-western community : I. *J. Educ. Psychol.*, 35, 357-368.

Hebb, D. O. (1949). *Organization of Behavior*. New York : John Wiley and Sons.

Hebb, D. O. (1958). *A Textbook of Psychology* Philadelphia : W. B. Saunders Co.

Hofstaetter, P. R. (1954). The changing composition of ' intelligence ' : a study in T technique. *J. Genet. Psychol.*, 85, 159-164.

Hunt, J. McV. (1961). *Intelligence and Experience*. New York : Ronald Press.

Kaiser, H. F. (1960) The application of electronic computers to factor analysis. *Educ and Psychol. Meas.*, 20, 141-151.

Lorge, I. (1952). ' Intelligence and cultural differences ' as seen by the test-makers. *Teach. Coll. Rec.* 54, 190-193.

Love, M. I., and Beach, S. (1957). Performance of children on the Davis-Eells Games and other measures of ability. *J. Consult. Psychol.*, 21, 29-32

MacArthur, R. S. (1962). The construct validity of a number of measures of intellectual potential for an Alberta Metis sample. Paper presented at Annual Meeting of Canadian Psychological Association.

Marquart, D. I., and Bailey, L. L. (1955) An avaluation of the Culture-Free Intelligence Test. *J. Genet. Psychol.*, 86, 353-358.

Nisbet, J D. (1953). Family environment : a direct effect of family size on intelligence. Occasional paper on Eugenics London : Cassell and Co.

Piaget, J. (1947). *Psychology of Intelligence*. Translated by Piercy, M., and Berlyne, D.E. London : Routledge and Kegan Paul.

Smith, T. W. (1956). Comparison of test bias in the Davis-Eells Games and the CTMM. *Calif J. Educ. Res.*, 7, 159-163.

Stern, W. (1914). *The Psychological Methods of Testing Intelligence*. Baltimore : Warwick and York.

Stillwell, L. (1958). A study of the Davis-Eells Games. Unpublished Master's Thesis. Ohio : Kent State University.

Stoddard, G. D., and Wellman, B. L. (1940). Environment and the I Q. *N.S.S.E. Yearbook*, 39, (I), 405-442.

Thurstone, L. L. (1938). Primary Mental Abilities. *Psych. Monog. No.* 1. Chicago : University of Chicago Press.

Tyler, F. T. (1953). Comments on the correlational analysis reported in " Intelligence and Cultural Differences." *J. Educ. Psychol.*, 44, 288-295.

Vernon, P. E. (1950). *The Structure of Human Abilities*. London : Methuen and Co

Vernon, P. E. (1955). 'The Assessment of Children ' in The Bearings of Recent Advances in Psychology on Educational Problems. London : University of London Institute of Education

Vernon, P. E. (1960). *Intelligence and Attainment Tests*. London : University of London Press.

Wheeler, L. R. (1942). A comparative study of the intelligence of East Tennessee Mountain children. *J. Educ. Psychol.*, 33, 321-334.

Woodrow, H. (1938). The relation between abilities and improvement with practice. *J. Educ. Psychol.*, 29, 215-230.

Wrigley, C. (1958). Objectivity in factor analysis. *Educ. and Psychol. Meas*, 18, 463-476.

(*Manuscript received, October,* 1962.)

THE RELATION OF PSYCHOLOGICAL TEST SCORES TO ACADEMIC BIAS

By LIAM HUDSON

(The Psychological Laboratory, University of Cambridge)

SUMMARY. On entering the Sixth Form, English schoolboys choose between Arts and Science subjects. Effectively, this choice commits them to one career field or the other. The present research is aimed to discover whether psychological tests reflect the academic bias of such boys towards Arts or Science subjects ; and whether they give any indication of boys' aptitude for specialised work at university and in research. A battery of conventional psychological tests was given to a sample of 315 15 and 16-year-old schoolboys, drawn from five schools These boys had just finished the G.C.E. ' O ' level examination, and stood a good chance of reaching university.

Results show that the tests used do reflect academic bias surprisingly well. On the other hand, there are signs that the relation of tests scores to aptitude for university and research work will prove to be complex. Some of the most gifted young Science specialists, for example, have scores which bear no relation to the trend of results in the sample as a whole. The implications of these and other findings are discussed.

I.—INTRODUCTION.

On entering the Sixth Form, at the age of 15 or 16, the clever English schoolboy has to choose between Arts and Science subjects. The choice he makes determines to a considerable extent the eventual course of his career. If he becomes, say, a History specialist, he will find it difficult later on to change to Classics or French ; and even more difficult, if not impossible, to change to Mathematics or Biology. Obviously, the more relevant information he has before him when making this decision, the better.

It is this choice which forms the context for the present research. Its ultimate aim is a series of tests which would show a 16-year-old (*a*) roughly what his own potentialities are, and (*b*) what demands would be made on him by various careers at university and in research. In order to develop and validate such tests, a series of longitudinal experiments are needed, in which individuals' test scores as schoolboys are related to their achievements as adults. The present paper is merely an interim report on the first stage of one such project. It concerns the basic question : do test scores reflect a schoolboy's bias of ability towards Arts or Science at the age when the crucial choice is being made ? Although this relation of test scores to academic bias will be discussed in detail, no attempt will be made to show that the relation is stable through time, or more generally that the tests have predictive validity. The criteria used are those appropriate to the Lower Sixth Form, and these are not grounds for vocational advice.

Relatively little is known about the relation of test scores to academic bias. During the last twenty years, many attempts have been made to validate differential aptitude batteries, but these have met with little success (Super, 1957 ; Horst, 1959 ; Anastasi, 1961). A notable exception is Crawford and Burnham's (1946) report that the Yale Differential Aptitude Battery has differential validity in predicting college grades. They also report (p. 242) that certain undergraduates have marked and persistent verbal/spatial biases of test score. Similar biases have been reported by Bradford (1948) among British 11-plus candidates. The results in both cases imply that such biases of test score are related to biases of academic ability and interest; but these relations have not been substantiated (Burt, 1950).

II.—METHOD.

(1) *The Sample.*

The present results are based on a sample of 315 15 and 16-year-old schoolboys. These were drawn from five schools: two public boarding, one public day, and two grammar. All five have high academic reputations. In each case, the boys were tested as they entered the Sixth Form, having just taken the G.C.E. ' O ' level examination. They were selected as having a reasonable chance of reaching university. The general level of their ability was high, the mean intelligence test score being only a few points lower than that found among Cambridge undergraduates (Hudson, 1960).

(2) *Subject-Groups and Criterion-Groups.*

In analysing the results, two systems of classification are used: Subject-Groups and Criterion-Groups. In the first of these, boys are grouped in terms of their special subject: Classics, Modern Languages, History, Physical Science, or Biology. This method is straightforward and yields a number of interesting comparisons. On the other hand, it gives little indication either of a boy's aptitude for his special subject, or of his academic bias. The most useful analysis is one which relates test scores to criteria *both* of academic success *and* of academic bias. In the present case, this is done by using Criterion-Groups.

Both the academic bias and the academic success of each boy were assessed from questionnaires, staff reports and G.C.E. ' O ' level results. Four Criterion-Groups were then selected, with twenty-five boys in each: ' Arts Specialists ' consisting of boys who are highly successful in Arts subjects and poor at Science; ' Arts All-rounders " who are highly successful at Arts and adequate at Science; ' Science-All-rounders ' who are highly successful at Science and adequate at Arts; and ' Science Specialists ' who are highly successful at Science and poor at Arts. As they were restricted to boys who are highly successful in their own subject, these Criterion-Groups reflect both differences in academic bias and differences in academic level. Differences in academic *bias* are reflected in comparisons between the Criterion-Groups; differences in academic *success* are reflected in comparisons between the Criterion-Groups and the remaining 215 boys in the sample. Four Criterion-Groups rather than five were used because genuine all-rounders proved relatively hard to find. The number of boys in each of the Subject- and Criterion-Groups is shown in Table 1. (It is important to note that the two systems of classification are independent, in the sense that a Classicist may be found in either of the Arts Criterion-Groups or in Rest-of-Arts, but not, of course, among the Scientists.)

(3) *The Indices Used.*

General Intelligence. The test used was A.H.5, a group test of high-grade intelligence specially suited to student samples (Heim, 1956). The test is administered in two Parts: Part I being verbal and numerical, Part II being diagrammatic.

Accuracy. A number of accuracy scores were derived from A.H.5, the most useful being ' Worse Accuracy ' and ' Range Accuracy.' A boy's accuracy on Parts I and II of A.H.5 were calculated separately—the number of correct answers as a percentage of the number attempted. The lower of the two percentages was then defined as his Worse Accuracy, and the difference between the two as his Range Accuracy. The accuracy score given in Tables 2-5 is Worse Accuracy.

Spatial Ability. An experimental spatial test was used as none of those already published was sufficiently difficult (Heim and Watts, 1957). It was aimed to tax students' ability at manipulating geometrical shapes in their mind's eye, but in practice was found to correlate substantially with General Intelligence (r=approximately ·5).

Vocabulary. An experimental, multiple-choice form was used, published tests being too easy (Heim and Watts, 1961).

General Knowledge. An experimental open-ended test was used, involving factual knowledge about 100 men, famous in current affairs, literature, painting, music, science.

Cognitive Bias. A differential score expressed the bias of each boy's ability in the cognitive tests. This contrasted his Verbal ability on the one hand with his Numerical and Spatial abilities on the other. (These scores were derived from A.H.5, A.H.5 and the Spatial Test, respectively.)

Bias of Interests. Interests were assessed by a system of cross-checks. Five sources of information were used, although all five were not available in every case; questionnaire, form-ratings, general knowledge, school staffs, interviews. A differential score was used to measure each boy's bias towards ' cultural and political ' as opposed to ' practical and outdoor ' interests. A working definition of this second category of interest was reached with some difficulty. Certain activities contribute heavily to an Arts/Science discrimination: car and radio mechanics, model building, carpentry, photography, camping, rock climbing, sailing, bird-watching. While others, ostensibly similar, show a significant tendency to obscure it: cine-photography, model railways; hunting, shooting, fishing, cycling, skiing, foreign travel. Similar difficulties were found in the ' cultural ' sphere; interests in literature and painting contribute heavily to an Arts/Science discrimination, whereas an interest in music tends to obscure it. As the present aim is primarily discriminative, an operational definition of these categories was felt to be in order. So those interests already mentioned as obscuring the Arts/Science discrimination were left out of account.

With one exception, the seven indices mentioned above are reasonably independent of one another, the correlations between them ranging from zero to ·5. The exception concerns the scale of Cognitive Bias and the Spatial Test from which it is partially derived, r in this case being of the order of ·7. Scores on the first five of these indices are graded on a five-point scale, A-E. Grade A represents the top 10 per cent. of each school sample; B, the next 20 per cent.; C, the middle 40 per cent.; D, the next 20 per cent.; and E, the bottom 10 per cent. A fuller account of tests and methods is given in Hudson (1961).

III.—Results.

(1) *Subject-Groups.*

The first aim in presenting the results is to give test ' profiles ' for each Subject-Group; to show, for example, whether there are characteristic differences between Classicists and Historians. In the interests of clarity, Table 2 summarises these profiles qualitatively. The over-simplification which this involves is relatively unimportant, because the difference between any given pair of Subject-Groups is expressed quantitatively in Table 3. Table 3 indicates which of the differences shown in Table 2 are of real discriminative importance.

Table 2 shows that Classicists are typically of average ability or better, and that they are particularly strong in accuracy and vocabulary. Their interests are diverse and they have no marked cognitive bias. Modern Linguists are unremarkable, apart from a verbal cognitive bias and, very surprisingly, a *weakness* in vocabulary—an inexplicably large number of highly successful Linguists had vocabulary grades of D or E. Historians are clearly distinguished by low cognitive scores and a strong ' cultural and political ' bias of interests. Physical Scientists have high scores on cognitive tests, and a marked bias both of cognitive ability and of interest. Biologists are undistinguished except for a strong ' practical and outdoor ' bias of interests.

TABLE 1

THE SAMPLE TESTED, DIVIDED IN TERMS OF (i) SUBJECT-GROUPS, AND (ii) CRITERION-GROUPS.

	SUBJECT-GROUPS	n	CRITERION-GROUPS	n
ARTS	Classics	27	Specialists	25
	Modern Languages	39	All-Rounders	25
	History	68	Rest-of-Arts	84
		134		134
SCIENCE	Physical Science	142	Specialists	25
	Biology	39	All-Rounders	25
			Rest-of-Science	131
		181		181

TABLE 2

SUBJECT-GROUPS: A QUALITATIVE SUMMARY OF TEST SCORES.

	GENERAL INTELLIGENCE	ACCURACY	SPATIAL ABILITY	VOCABULARY	GENERAL KNOWLEDGE	COGNITIVE BIAS	INTERESTS	n
CLASSICS	Fair	V. Good	Good	V. Good	Fair	None	Diverse	27
MODERN LANGUAGES	Fair	Fair	Fair	Weak	Fair	Verbal	Diverse	39
HISTORY	Poor	Weak	Poor	Fair	Fair	Verbal	Cultural & Political	68
PHYSICAL SCIENCE	V. Good	Good	V. Good	Fair	Fair	Numerical & Spatial	Practical & Outdoor	142
BIOLOGY	Fair	Fair	Fair	Weak	Fair	None	Practical & Outdoor	39

TABLE 3

INDICES WHICH DISCRIMINATE BETWEEN ANY GIVEN PAIR OF SUBJECT-GROUPS.

	MODERN LANGUAGES	HISTORY	PHYSICAL SCIENCE	BIOLOGY
CLASSICS	Vocabulary	Spatial Accuracy	Interests	Interests Vocabulary
MODERN LANGUAGES ..		Intelligence	Cognitive Bias	—
HISTORY			INTERESTS Cognitive Bias Spatial Intelligence	INTERESTS
PHYSICAL SCIENCE				—

STRONG DISCRIMINATION=0%—15% errors.
MODERATE DISCRIMINATION=16%—25% errors.
Weak Discrimination=26%—35% errors.

In describing the discriminations between the various Subject-Groups, three standards of discriminative efficiency are established : Strong, Moderate and Weak. A discrimination between two groups is defined as Strong if the proportion of errors involved is 15 per cent. or less, as Moderate if the proportion of errors lies between 16 per cent. and 25 per cent., and as Weak if it lies between 26 per cent. and 35 per cent. For example : if 15 per cent. of Historians lie above a given cut-off score on the vocabulary test while only 25 per cent. of Classicists lie below it, the proportion of errors involved in discriminating between these two groups is 20 per cent., and the discrimination is defined as Moderate. These three categories are distinguished typographically in both Tables 3 and 5 ; and in each cell of these Tables, the indices are arranged in descending order of discriminative efficiency. The cut-off score used in each case is the median for the whole schoolboy sample. All the discriminations shown in Tables 3 and 5 are statistically significant ($P=0{\cdot}05$; X^2 Test with Yates's Correction).

Although none of the discriminations between Subject-Groups is Strong, Table 3 shows that there are a number of significant differences between boys who specialise in the various academic subjects. The most obvious is that separating Historians from the Physical Scientists. For the rest, Classicists are distinguished from Modern Linguists and Biologists by their high vocabulary and from both groups of Scientists by the bias of their interests. Modern Linquists are distinguished from Physical Scientists by their cognitive bias.

(2) *Criterion-Groups.*

The Criterion-Groups consist of boys who are highly academically successful ; and they differ only in their academic bias. So contrasts between Criterion-Groups should reveal whether academic *bias* is reflected in test scores ; and the contrast of the Criterion-Groups with the rest of the sample should reveal whether academic *success* is so reflected. The profiles of the four Criterion-Groups and the remainder of the sample are summarised qualitatively in Table 4.

In general, the Arts Specialists have poor cognitive scores. They are also distinguished from Scientists by their interests, and from all other Criterion-Groups by their cognitive bias. The other three Criterion-Groups tend to have high scores, and the differences between them are relatively slight. These lie in interests, cognitive bias and, perhaps, spatial ability. The two 'Rest' groups score rather similarly. It seems that they occupy middle ground between the Arts Specialists at one extreme and the remaining three Criterion-Groups at the other.

The efficiency of each index in distinguishing between any given pair of Criterion-Groups is shown in Table 5. Again, only statistically significant differences are mentioned. The results may be summarised as follows :

(*a*) On the whole, discriminations between Criterion-Groups are much stronger than those between Subject-Groups.

(*b*) Among *successful* schoolboys, one distinction is of over-riding importance : that which separates Arts Specialists (i.e., those who are poor at Science) from the rest.

(*c*) Relative to this gulf, the dichotomy between Arts and Science is less important. Only his interests reveal the direction in which a successful all-rounder has specialised.

(*d*) Academic bias is reflected excellently among Arts boys, but poorly among Scientists. Or—to express the same relation in reverse—the tests reflect weakness in Science excellently, but only poorly reflect weakness in Arts.

(*e*) The relative similarity of Arts Specialists and Rest-of-Arts groups, and the dissimilarity of both from Arts-All-rounders emphasizes that tests do not measure the qualities underlying success in Arts, and also that the Arts All-rounder is a relatively unusual type.

(*f*) There is a surprising, although non-significant, tendency for Arts Specialists' cognitive scores to be *worse* than those of the Rest-of-Arts group.

(*g*) The indices : General Intelligence, Accuracy, and Spatial Ability all distinguish boys who are good at Science from those who are poor. Nevertheless, very high Accuracy scores are largely restricted to *Arts* All-rounders, while the Accuracy scores of successful Scientists tend to cluster in the B and C grades. Vocabulary measures academic aptitude in Arts and Science alike ; and, to a lesser extent, the same seems to be true of General Knowledge. Cognitive Bias gives a good indication of academic bias, particularly among Arts boys ; and Bias of Interests distinguishes Arts boys as a whole from Scientists.

Tables 4 and 5 do not give a clear idea of how efficiently the various academic groups could be distinguished in practice, if all seven indices were combined. As a rough check on the strength of these discriminations, the scores of all the Criterion-Group members were placed on cards and an attempt was made to identify each boy in the light of the evidence given in Tables 4 and 5. The results, shown in Table 6 are most satisfactory :

(*a*) Two-thirds of the Criterion-Group members are identified correctly, whereas only some 25 per cent. would have been indentified correctly had they been sorted at random. As many as 97 out of 100 are placed either correctly or in a neighbouring group.

TABLE 4

CRITERION-GROUPS : QUALITATIVE SUMMARY OF TEST SCORES.

	General Intelligence	Accuracy	Spatial Ability	Vocabulary	General Knowledge	Cognitive Bias	Bias of Interests	n
Arts Specialists	Poor	Poor	Poor	Good	Fair	Verbal	Cultural & Political	25
Arts All-Rounders ..	V. Good	V. Good	Fair	V. Good	Good	None	Cultural & Political	25
Science All-Rounders	V. Good	Good	Good	V. Good	V. Good	None	No Bias	25
Science Specialists ..	V. Good	Good	V. Good	Good	Good	Numerical & Spatial	Practical & Outdoor	25
Rest-of-Arts	Weak	Weak	Weak	Fair	Weak	None	No Bias	82
Rest-of-Science	Fair	Fair	Fair	Weak	Fair	None	Practical & Outdoor	133

TABLE 5

INDICES WHICH DISCRIMINATE BETWEEN ANY GIVEN PAIR OF CRITERION-GROUPS.

	Arts All-Rounders	Science All-Rounders	Science Specialists	Rest-of-Arts	Rest-of-Science
Arts Specialists ..	ACCURACY COGNITIVE BIAS INTELLIGENCE Spatial	SPATIAL ACCURACY INTELLIGENCE Cognitive Bias Interests	*SPATIAL* *COGNITIVE BIAS* INTERESTS INTELLIGENCE Accuracy	—	INTERESTS COGNITIVE BIAS Spatial Accuracy Vocabulary
Arts All-Rounders		INTERESTS	INTERESTS	INTELLIGENCE Accuracy Vocabulary Cognitive Bias	*INTERESTS* VOCABULARY Intelligence
Science All-Rounders			Cognitive Bias	INTELLIGENCE GENERAL KNOWLEDGE Vocabulary Accuracy Spatial	VOCABULARY General Knowledge Intelligence
Science Specialists ..				SPATIAL COGNITIVE BIAS INTELLIGENCE Interests Accuracy	Intelligence Spatial
Rest-of-Arts					Cognitive Bias Interests

STRONG DISCRIMINATION= 0%—15% errors.
MODERATE DISCRIMINATION=16%—25% errors.
Weak Discrimination=27%—35% errors.

TABLE 6

An Attempt to Identify Criterion-Group Members, Using All Seven Indices Together.

	Identified as— " Arts Spec."	"Arts All-R "	" Sc. All-R "	" Sc. Spec."
	n	n	n	n
Arts Specialists	18	6	1	—
Arts All-Rounders	3	19	2	1
Science All-Rounders	1	5	14	5
Science Specialists	—	—	9	16

(*b*) There is some confusion between neighbouring groups, particularly between the two groups of Scientists ; but the discrimination between the other groups is good, and that between the two extremes, the Arts and Science Specialists, is almost complete. Not one Arts Specialist was identified as a Science Specialist, or *vice versa*.

(*c*) The discrimination between Arts and Sciences is also surprisingly strong ; only four Arts boys were identified as Scientists, and only six Scientists were identified as Arts boys.

The samples used here are rather small, and a certain degree of chance error is inevitable. When the research is extended to other schools, particularly those with different academic traditions, changes must be expected. (It has been found, for instance, that the intellectual calibre of Biology specialists varies considerably from one school to another.) It seems unlikely, on the other hand, that the major differentiations will be materially altered.

(3) *Other Findings.*

(i) *Gifted Scientists.* The general level of test scores among successful Physical Science specialists is high ; and, in practice, most gifted young scientists are easily identifiable from their scores. There are, nevertheless, a few exceptionally gifted Physical Science specialists whose scores are unaccountably low. Perhaps three cases are sufficient to illustrate this type : D., a mathematical ' prodigy ' ; P., a Chemistry specialist who has since gained an Open Scholarship at Cambridge ; and R., the best of a remarkable group of young scientists at the public day school. Their test scores are respectively : General Intelligence—C, D, C ; Accuracy—E, C, D ; Spatial Ability—A, D, C ; Vocabulary—B, B, C ; General Knowledge—C, B, A. All three have strong or fairly strong interests, but none has either a ' political and cultural ' or ' practical and outdoor ' bias. All three are considered by their form-mates to be hard-working. In general, their remarkable talents reflect themselves in their test scores in hardly a single respect.

Although such cases account foı a mere 2–3 per cent. of the whole sample, they demand an explanation. The most obvious is that the boys concerned were not trying. With most of them, this can be ruled out of account. Most of them are notably conscientious ; and as their academic reputation usually preceded hem, they were watched with special care during the testing sessions. A second possibility is that these boys were made over-anxious by the novelty of the testing situation. Again, this seems feasible in a few cases but not in all. There are marked differences of personality and emotional maturity within this group, so it is unlikely that a single emotional factor could account for all of them. A third and much more likely explanation is that there are important scientific abilities which the present tests miss. There is, after all, no *prima facie* reason for believing that the complexity of scientific work can be epitomised in the present handful of paper-and-pencil puzzles. Indeed, it would be strange if it could. There are several normal scientific skills which conventional tests take little account of; the 'inductive' skills of marshalling complex evidence, for example, and accurate generalisation from incomplete evidence. Also there is the whole area of creativity (e.g., Getzels and Jackson, 1962). If new testing methods are needed, the best way of discovering them is presumably a detailed study of present anomalies.

(ii) *Extreme Arts Specialists.* Another tendency obscured in the general presentation of results concerns a small group of Arts Specialists. These are boys who are outstanding at Arts subjects but very weak at subjects which tax their powers of formal reasoning, e.g., Mathematics, Physical Science, Latin. This group shows a strong *negative* relation between such measures as General Intelligence, Accuracy, Spatial Ability on the one hand, and General Knowledge on the other. The scores of three of these extreme Arts Specialists—C., G., S.—are: General Intelligence—E, E, E; Accuracy—E, E, E; Spatial Ability—E, E, E; Vocabulary—B, C, B; General Knowledge—A, A, A. All three have a very strong 'political and cultural' bias of interests, and a heavily verbal bias of cognitive ability. Both C. and G. are exceptionally hard-working. The General Knowledge scores of these three boys are, respectively, the fourth highest, third highest and the highest in the whole sample; while, with one exception, G. has the lowest General Intelligence in the whole sample, and C. the lowest but one. As with the Physical Scientists mentioned above, this type is exceptional; such boys account for less than 5 per cent. of the whole sample. Yet, they are possibly of even greater practical importance, because it is precisely this type of boy who is filtered out of the academic system by university requirements in Latin and Mathematics.

(iii) *Negative and Compensatory Relations.* This negative relation among extreme Arts Specialists seems to be one instance of a more general pattern. Among Historians and Modern Linguists there is a marked tendency for a high score on any one test to be counter-balanced by low scores on others. In cases where a boy's scores were all low, he was found almost invariably to be exceptionally hard-working. So it seems that in History and Modern Languages certain mental abilities are mutually inter-changeable; and that where ability is lacking, hard work will serve in its stead. It will be interesting to see which of the various test profiles found among Arts boys bode well academically; and whether qualitative differences of intellectual type are detectable between them. Why this form of equivalence should not hold among Classicists, Physical Scientists or Biologists is not clear. One possibility is that History and Modern Languages admit a wide range of differences in intellectual 'style,' whereas in Classics and Science there may be a greater tendency for one intellectual method to be the best.

These results suggest that a compensatory mechanism may be at work in some of the boys concerned; that special weakness in, say, formal reasoning has caused them to concentrate on general knowledge. Compensation of this kind is illustrated in a case of sibling rivalry. These two brothers were 18 months apart in age, yet the younger had overtaken the older academically. They had an unsatisfactory home background ; and at the time of his interview, the elder brother showed obvious signs of emotional disturbance. The younger brother was a Physical Scientist with cognitive scores above the average, a spatial bias of abilities, poor general knowledge and exclusively practical interests. The elder was a Historian with very low cognitive scores, a verbal bias of ability, good general knowledge, and exclusively cultural interests. In effect, the elder boy had concentrated his efforts in precisely those areas where his brother was weak.

(iv) *Accuracy, Emotional Stability and Originality.* Previous work has shown that accuracy scores can be used as an index of emotional stability (Petrie, 1948). In the present research, both Worse and Range Accuracy were found to correlate substantially with interview ratings of emotional stability (Hudson 1961)—r=·42 and ·65, respectively ; n=51. Individual cases also suggested that there may be a connection between accuracy scores and potential originality. The Worse Accuracy scores of the six exceptionally gifted boys mentioned above were poor. The same is true of their Range Accuracy, their grades being E, C, D and E, E, D for Arts Specialists and Physical Scientists, respectively. Among Physical Scientists such poor accuracy scores are relatively rare, and generally are a sign of academic mediocrity. Although superficially anomalous, this finding is reminiscent of some previous suggestions about the nature of originality. Guildford (1950), for example, has suggested that creativity may stem from a general distaste for repetition ; and Getzels and Jackson (1962) found ' creative ' students impatient with the routine of classroom work.

IV.—CONCLUSION.

With certain important reservations, we may conclude that the academic bias of clever schoolboys does reflect itself in test scores with surprising clarity. If this relation holds true at university and in research, the present tests could be of some use in advising schoolboys about their choice of career. On the other hand, there are good reasons for expecting this relation to weaken. The demands made by a particular subject, Physics for example, may well change radically between G.C.E. ' O ' level and Tripos ; work for Tripos may differ from research ; and the various branches of research in an area like Physics may differ among themselves (Roe, 1953). So although conventional tests may be said to have proved their worth at the Sixth Form level, the influence of ill-defined qualities like originality and maturity of judgment may well increase as the present sample matures ; and if they increase, conventional tests may tend to lose their value.

Even at the Sixth form level, much remains to be done. The present tests do not reflect ability in Arts subjects, nor do they distinguish certain exceptionally gifted young scientists. New tests are certainly needed, but there is little indication of what shape they should take. The abilities which the present tests fail to detect may be primarily cognitive, e.g., the ability to marshall complex evidence. They may be more personal, e.g., imaginativeness, maturity of judgment, insight into human motives. Or they may involve differences of intellectual attitude and ' style ' (Rapaport 1957 ; Getzels and Jackson 1962).

The fact remains that there are brilliant boys with relatively poor test scores, and seemingly dull boys whose scores are relatively high. As Getzels and Jackson point out, conventional psychological tests normally account for less than a quarter of the variance in academic success ; and it seems that the deductive, puzzle-solving abilities which these tests do measure have little bearing on intellectual originality. So although there is scope for new cognitive tests of the orthodox kind, the most fruitful research may well lie in those areas of attitude and 'style' where cognitive and personal qualities interact. In the past there has been a tendency to approach these areas obliquely, using projective techniques like the Rorschach (Rapaport 1957.); but distinctions of type emerging from this work have not yet been related to differences of attitude and 'style' discernible in academic practice. Obviously, if these distinctions are to stand any chance of explaining intrinsically interesting phenomena, this gulf between the laboratory and the world at large will have to be bridged. It is at this point that the skills of the natural historian become vital. If test scores are to be related to the real world—and this, after all, is what validation should consist in—mental testers are obliged to observe intellectual workers at work, to analyse precisely what they do and how they do it. The immediate task, in other words, is to find ways of describing and quantifying those intellectual qualities which do not lend themselves to quantification automatically. A good example of the exploratory work which this will involve is Getzels and Jackson's (1962) study contrasting 'highly intelligent' and 'highly creative' school children. This study not only shows sensitivity to *qualitative* differences in school children's thinking, but also tackles the formidable problems of description which these qualitative differences present. Perhaps it is a good omen for the future of mental testing that the statistics these authors use are largely rudimentary.

Finally, the limitations of the present results must be reasserted. In the first place, the intellectual demands made by the various Sixth Form subjects are relatively elementary; in the second place, the choice between them by no means determines the course of a schoolboy's career. A Sixth Form History specialist may end up as an accountant, businessman, politician, anthropologist, soldier, barrister, clergyman. He may have talents unrelated to his job—he may work in a bank and write poetry, or in a customs house and paint. Or he may have talents which circumstances force him to neglect. The present categories are bound to be provisional, and no really satisfactory assessment can be made until the present sample has been followed through university and beyond. Their test scores as schoolboys can then be related not merely to formal criteria like degree class, but to full descriptions of their intellectual lives as adults.

ACKNOWLEDGEMENTS.—I am deeply indebted to the headmasters and members of staff whose co-operation has made this work possible; to Professor O. L. Zangwill, Professor P. E. Vernon and Dr. A. W. Heim for their help and advice; to Dr. A. W. Heim for making available unpublished testing material; and to the Ministry of Education and Department for Scientific and Industrial Research who have given the work their financial support.

V.—REFERENCES.

ANASTASI, A. (1961). *Psychological Testing*, 2nd edit. New York: Macmillan.

BRADFORD, E. J. G (1948). Symposium on the selection of pupils for different types of secondary schools. *Brit. J. Educ. Psychol*, 18, 67.

BURT, C. (1950). Symposium on the selection of pupils for different types of secondary schools. *Brit. J. Educ. Psychol*, 20, 1.

CRAWFORD, A. B, and BURNHAM, P. S. (1946). *Forecasting College Achievement*. Yale: University Press.

GETZELS, J. W., and JACKSON, P. W. (1962). *Creativity and Intelligence*. New York: Wiley.

GUILDFORD, J. P. (1950). Creativity, *American Psychologist*, *5*, 444.

HEIM, A. W. (1956). *Manual for the Group Test of High Grade Intelligence A.H.5*. London: National Foundation for Educational Research.

HEIM, A. W, and WATTS, K. P. (1957). An experiment on practice, coaching and discussion of errors in mental testing. *Brit. J. Educ. Psychol.*, 27, 199.

HEIM, A. W, and WATTS, K. P. (1961). A preliminary study of the self-judging vocabulary scale. *Brit. J. Psychol*, 52, 175.

HORST, P. (1959). Differential prediction of academic success. *Report*, Office of Naval Research Contract 477 (08) : Univ. of Washington.

HUDSON, L. (1960). A differential test of arts/science aptitude. *Nature*. London, 186, 413.

HUDSON, L. (1961). Arts/science specialisation : a study of differential aptitude. *Unpublished Ph.D. Thesis* : Cambridge University.

PETRIE, A. (1948). The selection of medical students. *Lancet*, 2, 325.

RAPAPORT, D. (1957). " Cognitive Structures." In *Cognition : the Colorado Symposium*. Cambridge, Mass. · Harvard Univ. Press.

ROE, A. (1953). A psychological study of eminent psychologists and anthropologists, and a comparison with biological and physical scientists. *Psychol. Monogr.*, 67, No. 352.

SUPER, D. E. (1957). The Multifactor tests : summing up. *Pers. Guid. J.*, 36, 17.

(*Manuscript, received, 27th October*, 1962.)

CHANGES IN TERMAN-MERRILL I.Qs. OF EDUCATIONALLY SUB-NORMAL BOYS

BY

C. S. RUSHTON

(Sometime Educational Psychologist, Kent School Psychological Service)

AND

A. E. STOCKWIN

(Sometime Deputy Headmaster, Seabrook Lodge School, Hythe)

SUMMARY. Test results on the Terman-Merrill* 1937 revision of the Stanford-Binet Intelligence Scale (Form L) were examined for 111 'educationally subnormal' boys in a residential special school. A comparison between results on entry at the age of 7 years and above, and before leaving between 14 and 16 years, showed a deterioration in mean I.Q. scores, despite the fact that the Roberts-Mellone† (1952) corrections were applied.

The proportion of 'verbal subtests' passed was consistently lower than the proportion of 'non-verbal' tests,' at all age levels.

The hypothesis that the deterioration in mean I.Q. scores between mean C.A. 9¼ and 15¼ years, was due partially to the increased verbal weighting on this test at the higher age levels, was examined and supported.

It was suggested that the test under consideration is unsuitable for the all-round cognitive assessment of educationally sub-normal‡ children. Suggestions were made for a more suitable type of test, and for further research on Terman-Merrill (L).

I.—INTRODUCTION AND PREVIOUS RESEARCH.

PREVIOUS investigations on this subject have yielded conflicting results. Many of the earlier findings were consistent in that the mean I.Q. of dull children appeared to decline with age. Thus, Phillips (1940) concluded that for children with Terman-Merrill I.Qs. between 50 and 80, there was a tendency for a general and substantial decline which increased progressively with age. Parker (1934), Thorndike (1940) and Stott, (1960) all show increasing deterioration of I.Q. scores with increasing interval between test and retest.

Perhaps the first and most striking contradiction of these findings came from Wellman (1934), who reported that all the groups which she tested tended to show improved I.Qs. while 'the most backward group' showed the greatest increase. Schmidt (1946) and the Clarkes (1958 and 1959) also reported striking gains in the I.Qs. of the 'feeble-minded,' but the former study (Schmidt's) was severely criticized by Kirk (1948) and cannot now be taken seriously.

More recent studies have had available more accurate measuring techniques. Collman and Newlyn (1958) for instance, on a test-retest study of educationally subnormal pupils in special schools in Essex, found that the use of the R.M. correction table (1952) was fully justified. On re-test they found no significant mean change in I.Q.—but this was after a period of only one year. Hiskey and Sadnavitch (1958) and Scarr (1953) report similar findings.

---

* From this point onwards referred to as T-M (L).

† From this point onwards referred to as R.M. corrections.

‡ From this point onwards referred to as E.S.N. children.

In a private communication, P. F. Portwood states that when I.Q. scores on T-M (L) were considered without the R.M. corrections being applied, there appeared to be a significant drop in I.Q. scores after a two-year interval. When the corrections were applied, however, the significant drop (due to the increasing S.Ds. between these age levels and thus to a spurious statistical effect) was changed to a slight though statistically insignificant increase. Here again is evidence which appears to confirm empirically the validity of the R.M. corrections, and the advisability of using them when dealing with T-M (L) I.Qs. on E.S.N. pupils.

II.—PRESENT INVESTIGATION—METHOD.

A comparison was made between the T-M (L) I.Qs. of 111 E.S.N. boys in a residential special school in Kent on ' First Testing ' (usually at the age of 8 or 9 years), and ' Last Testing ' (between the ages of 14 and 16 years). All testing was done by assistant school medical officers or by the deputy headmaster, who had been trained in the administratisn of T-M (L).

For testings below C.A. 15 years, R.M. corrections were applied to all I.Qs. Above the age of 14 years 11 months no R M. corrections are published, yet for the 15-year-old age group the mean I.Q. is 102·8 and S.D. 19 (Terman and Merrill 1937), and thus needs correction. MacNemar, in his earlier correction table (1942) used the same corrections for C.As. 14y. 6m. to 15y. 5m., as Roberts Mellone do for C.As. 14y. 3m. to 14y. 5m. (1952). Thus, it was considered justifiable to use these corrections for test results in the present investigation between C.A. 15y. 0m. and 15y. 5m. Above this age MacNemar does not offer a correction, and at the 16-year age-level S.D. drops to 16·5 (Terman and Merrill, 1937). Thus, for all test results above C.A. 15y. 5m., no correction was made.

Means and S.Ds. were computed on these I.Qs., and significances of differences between means were obtained, using Sandler's ' A Test ' (Sandler 1955), which he claims has been ' rigorously derived ' from Students ' *t* ' Test, and corresponds to the two-tail form of ' *t*.'

III.—RESULTS AND THEIR TREATMENT.

A. *Subjects.* Of the 111 boys studied :

(1) 69 per cent. came from homes in ' urban ' areas, and 31 per cent. from ' rural ' areas.

(2) 71 per cent. came from families where the father was ' unskilled ' or ' semi-skilled ' (Grades IV and V of the Registrar-General's categories of occupation groups), as contrasted with 26½ per cent. in the general population of Kent (H.M.S.O. 1954). This finding is in line with the results of earlier researches. (Jones 1954).

(3) Only 29 per cent. came from families in the ' higher ' occupation groups (Registrar-General Grades I, II and III), as contrasted with 73·4 per cent. in the general Kent population (H.M.S.O. 1954).

(4) 62 per cent. came from families where there were three or more children, as contrasted with the normal 27 per cent. of families in Kent of this size (H.M.S.O. 1954). On the other hand, only 14 per cent. of our E S.N. population were ' only ' children, as opposed to 42 per cent. in the general Kent population (H.M.S.O. 1954). Eysenck (1953) also reports that dull children tend to come from large families.

B. *Test Results.*

(1) For the population of E.S.N. boys tested, there is a statistically significant 'deterioration' in mean T-M(L) I.Qs. between the ages of 9y. 2m. (before entry to residential special school) and 15y. 6m. (just before leaving special school). (See Table 1).

TABLE 1

TEST-RETEST DATA T-M (L) I Qs. WITH R.M. CORRECTIONS APPLIED.

(N=111)	First Testing	Last Testing
I.Qs. : Means S Ds. Range	 71·0 7·91 47 to 96	 66·0 9·43 46 to 93
C.As. : Means S.Ds. Range	 9y 2m 1·72y. 6y. 11m. to 13y. 4m.	 15y. 5½m. 0·67y. 14y. 0m. to 16y. 1m.
Difference between I.Qs on first and last testing	5·0 I.Q. points	
Sandler's A	·0308	
Level of Significance	Below ·001	
Product Moment " r " of I.Qs.	+0·626	

TABLE 2

ANALYSIS OF I Q. CHANGES ON T-M (L) WITH R-M CORRECTIONS APPLIED.

	I.Q. Deteriorated more than 3 points		I.Q. Constant ±3 points		I.Q. Improved more than 3 points		Chi^2	Level of Significance
	N	%	N	%	N	%		
Total N=111	68	61%	27	24½%	16	14½%	4·41	P ·05
First I.Qs.	N		N		N			
Below 60 61 to 70	1 22		5 8		7 3		4·59	P ·05
Below 60 71 to 80	1 39		5 10		7 5		5·06	P ·05
Below 60 81+	1 6		5 4		7 1		3·17	P at ·07

(2) The correlation coefficient (r=+0·626) is in general agreement with the figure which one would expect after a test-retest interval of about 6½ years. Eysenck (1953) and Stott (1960) report that the evidence from thirty similar test-retest studies indicates a correlation of ·79 after 4 years, with a drop of about ·04 points in the coefficient for every additional year intervening. On these predictions, for a test-retest interval of 6½ years, one would expect 'r' to be +0·69., whereas we find +0·626. This lower correlation may well be due to the fact that many different school medical officers were involved in the testing.

(3) From Table 2, it is clear that many more individuals show 'deterioration' than 'improvement' in I.Q. on final testing. Chi^2 test, where 'I.Qs. constant' were apportioned equally to 'improved' and 'deteriorated,' showed that this greater number of deteriorating scores was significant at a ·05 level.

(4) If one sub-divides the data, however, into I.Q. ranges, it can be seen that there appears to be a 'cut-off point' around I.Q. 60 on first testing. Those boys with an initial I.Q. of below 60 on first testing, tend to show an improved I.Q. (more than 3 points) on final testing. Those boys with an I.Q. of more than 60 on first testing tend to show a deterioration of I.Q. on final testing of more than 3 I.Q. points.

These differences in the direction of I.Q. change appear to be statistically significant, and greater than one would expect by chance alone. It may be that they are partially explicable by the concept of 'statistical regression towards the mean" (McNemar 1940 and Clarke, Clarke and Reiman 1958), since they occur in different directions in the two extreme groups, and *towards* the group means. The fact that the group mean itself changes between first and last testing cannot, however, be explained on these grounds, as it changes in the 'wrong' direction, i.e., *away* from the mean score of the population.

IV.—EXPLANATORY HYPOTHESIS.

A.—*Procedure and Findings.*

In a private communication, Professor A. D. B. Clarke, suggested that the most likely reason for the decline in scores was that " the test material is more verbally biassed on retest (Mean C.A. 15·5½), than on initial test " (Mean C.A. 9y. 2m.). His argument is based on the hypothesis that " the increased verbal weighting at the higher age levels is responsible for decreasing the final I.Q., since commonly verbal assets are more seriously impaired in E.S.N. children than are performance assets." (See also Tansley and Gulliford 1960).

In order to test this hypothesis, the sub-tests of T-M (L) were classed as either 'Verbal' (V) or 'Non-Verbal' (including 'Arithmetical') (P) on the basis of the classification of test responses published as Appendix I in Kennedy-Fraser (1948). The original test sheets of fifty-one boys which were immediately available were then re-examined, and each sub-test attempted at either first or final testing was labelled 'V test' or 'P test.' Totals and Chi 2 of their differences are shown in Table 3.

B.—*Indications and discussion.*

(1) On Verbal Tests there is no statistically significant difference between success rates on first and final testings. The direction of the change, however, is towards an improvement in 'verbal efficiency.'

(2) On Non-Verbal Tests there is no statistically significant difference between success rates on first and final testings. The direction of the change, however, is towards a deterioration in 'non-verbal efficiency.'

TABLE 3

Differences Between Verbal and Non-verbal Tests on T-M (L) with R-M Corrections Applied.

(N Boys=51)

	Verbal Tests			
	Passed	Attempted	Percentage Passed	Chi^2
First Testing	352	927	38%	0·865 P=·40 NOT significant
Last Testing .. .	689	1580	43·6%	
	Non-Verbal Tests			
	Passed	Attempted	Percentage Passed	Chi^2
First Testing	349	519	67·3%	2·80 P=·10 NOT significant
Last Testing	192	328	58·5%	
	First Testing			
	Passed	Attempted	Percentage Passed	Chi^2
Verbal Tests . ..	352	927	38%	58·2 P=less than ·001 Highly Significant
Non-Verbal Tests	349	519	67·3%	
	Final Testing			
	Passed	Attempted	Percentage Passed	Chi^2
Verbal Tests ...	689	1580	43·6%	12·71 P=less than ·001 Highly Significant
Non-Verbal Tests	192	328	58·5%	

(3) On both first and final testing there is a consistent and highly significant difference between success rate on Non-Verbal tests and on Verbal Tests. E.S.N. boys do better on P tests than on V tests from ages 8 and 9 years up to ages 15 and 16 years. Generally speaking, they passed only four out of 10 V tests, but six out of ten P tests.

Is Clarke's hypothesis that " the increased verbal weighting at higher age levels is responsible for decreasing the final I.Q. " supported ?

The experimental data show that on first testing, out of a total of 1,446 sub-tests attempted, 64 per cent. were V tests, while only 36 per cent. were P tests. On final testing, however, out of a total of 1,908 sub-tests attempted, 83 per cent. were V tests, and only 17 per cent. P tests. Since there is no statistically significant difference on V tests between success rate on first testing and on final testing, let us presume that there is no difference at all. We may assume that on V tests the success rate is 40·8 per cent. (mean of 38 per cent. and 43·6 per cent., the actual success rates achieved). Similarly for P tests we assume a success rate of 62·9 per cent. (mean of 67·3 per cent. and 58·5 per cent., the actual success rates obtained).

Let us next presume that on *first* testing 100 tests were attempted—of these sixty-four would be V tests and thirty-six P tests. Applying the differential success rates just calculated, this would mean that twenty-six V tests would be passed, and 22·6 P tests would be passed. Total tests passed on first testing would thus be 48·6.

If we then presume that on final testing, 100 tests were attempted, of which eighty-three would be V tests and seventeen would be P tests, applying the differential success rates again, this would result in thirty-four V tests being passed, and 10·7 P tests. The total tests passed on final testing would, therefore, be 44·7 as opposed to 48·6 on first testing.

Thus, if the actual number of tests attempted on first and last testing was held constant at 100, then considering the different proportion of V and P tests attempted, there would, on balance, be an apparent deterioration in the numbers of sub-tests passed, given no change at all in ' efficiency level ' (i.e., in success-rate) from one testing to the other.

In this hypothetical case then, it would appear that Clarke's hypothesis is supported, i.e., that the increased verbal weighting on final testing is responsible for decreasing the final score.

If we now consider the actual ' raw scores ' from the experiment, we find the success rates for V tests and P tests combined, to be 48·5 per cent. on first testing, and 46·2 per cent. on final testing. Chi^2 of this difference is 0·93, which is not statistically significant.

These percentages are similar to the proportions passed in the hypothetical case. The fact that in the real experiment the percentage of all sub-tests passed on final testing is higher than that in the hypothetical case, must be due to the slightly improved ' verbal efficiency ' on final testing. (See Table 3). 83 per cent. of all the sub-tests on final testing were verbal and the improved ' verbal efficiency ' on these outweighs the influence of the deteriorated ' non-verbal efficiency,' on the remaining 17 per cent. of tests taken.

C.—*General discussion.*

This ' over-weighting ' of T-M (L) with verbal sub-tests from Yr. VII level and upwards, has been criticized elsewhere (Clarke and Clarke 1958 and *Bull. Brit. Psychol. Soc.* 1958). The results of the present investigation would tend to supoprt these criticisms, though not at an acceptable level of statistical significance. The fact that T-M (L) ' markedly penalizes ' the dull person, especially if he is from ' poor or adverse surroundings,' would make it an unsuitable measuring instrument for use in the testing of E.S.N. children, since although it ' indicates his verbal handicaps ' effectively, it fails to give us an all-round assessment of his wider intellectual capacities.

In the present study, Roberts-Mellone corrections were used throughout, a procedure which has been confirmed empirically since the publication of their original paper (1952), following MacNemar (1942), by Collman and Newlyn (1958), by Hiskey and Sadnavitch (1958), by Scarr (1953) and by Portwood (private communication). ·Such evidence would appear to outweigh the arguments of Stott (1960), which are not always clear, neither are his basic assumptions universally acceptable.

Some of Stott's findings (e.g., greater deterioration in the scores of girls than of boys) suggest that the consistent deterioration which he finds may be due partially " to real changes in the rate of mental development " of such children, yet it is notable that he concludes, " this is not to deny that part of the variation is also due to faulty standardization." Present results would suggest that the deterioration is partially due to the actual structure of the test, which varies in its composition from one mental age level to another, and tends to lower the scores of those children with comparatively poor verbal ability, in direct proportion to the verbal weighting in the selection of sub-tests used at any one testing.

Stott's (1960) paper fails to supply adequate information about the chronological ages of the children tested, about their mental ages, or about the average re-test interval for the children re-tested after 3 years 5 months. The present study offers data on a more homogeneous population, with a test re-test comparison between scores at junior school age, and then nearing school-leaving age. It is notable that in some of the studies discussed above, whether or not R.M. corrections were applied, there is a greater deterioration in test scores, the greater the interval between test and re-test. In the present study, where the test re-test interval is longer than in most, there is a greater deterioration, which suggests that the same aetiological factors may well be cumulatively operative on all the subnormal populations studied.

Various theoretical issues remain to be considered. The work of the Clarkes (1958a, 1958b and 1959), of Mundy (1957), and of Penrose (1949) strongly suggests that for the sub-group we have been studying, there is a relative constancy of ' mental efficiency ' (and thus of I.Q.), during childhood, but that after the age of about 15 years, there are often marked improvements in I.Q. Clarke's (1959) hypothesis is that such increments amongst the ' feeble-minded ' represent " not a response to present conditions, but rather a recovery from early deprivation," maternal or environmental. This work by no means contradicts the well-known thesis of Bowlby (1951 and 1956) on deprived children, whose follow-up studies have, unfortunately, as yet all been on children below the age of 14 years ; it merely has a bearing on the long-term outcome.

O'Connor (1956) points out that the concept of deprivation is still confused, but that one important variable in the normal development of abilities and character, is a ' stimulating environment.' This point of view is supported by Sontag, Baker and Nelson (1958). They found that amongst pre-school children ' emotional dependence on parents ' was associated with deteriorating I.Qs. during that period—such children were establishing behaviour patterns which were not conducive to ' learning to learn ' during these early years. On the other hand, pre-school children who developed behaviour patterns characterized by aggressiveness, self-reliance, initiative and competitiveness (those from ' stimulating environments '), were developing personality traits which result during school life in gradually improving I.Qs. Such children seem to develop a high need for achievement, which operates as a motive for learning.

In the present paper it is established that E.S.N. boys from ' working class homes ' (i.e., from the two lowest social-occupation groups, from which 71 per cent. of our population were derived), score consistently better on the non-verbal tests in T-M (L) than on the verbal tests. Bernstein (1958 and 1959) found that of 309 G.P.O. messenger boys in adolescence from ' unskilled and semi-skilled backgrounds,' 206 scored Matrices I.Qs. of 101+ while the means of their Mill Hill Vocabulary scores fell below this.

It has been argued by Hollingworth (1920) and others that in subnormal children the curve of ' mental development ' rises constantly, but more slowly than in normal, and that there is a cessation of growth at an earlier age than for normal children. If this is so, and the curve of mental development is 'less than linear,' then the I.Q. of subnormal children must inevitably fall. Some support has been given to this hypothesis by earlier investigators such as Phillips (1940). In more recent studies, however, where the varying S.Ds. of T-M (L) have been corrected for, no such deterioration has been found, e.g., Collman and Newlyn (1958), Hiskey and Sadnavitch (1958) and Portwood. We have seen too in the work of the Clarkes that marked rises in I.Q. level may occur at adolescence; and Davis (1957) has challenged the earlier assumption, arguing that environmental influences are extremely important in influencing the course of 'mental development.'

The influence of environment may also be important in determining the significant differences found in the present study between the efficiency of E.S.N. boys on non-verbal and verbal tests. We know that they tend to come from large 'working class' families, where parental training, encouragement and concern tend to be lacking, and where other investigators have found children with poor verbal and academic abilities (Nisbet 1953, Anastasi 1956). Such children, like adolescent delinquents (Glueck and Glueck 1950), tend to do relatively well on 'non-verbal tests' whether they are of a performance or perceptual type. Pringle and Bossio (1958) have presented evidence that groups of deprived children obtain lower Verbal than Performance W.I.S.C. I.Qs. Tansley and Gulliford (1960) report that E.S.N. children in a residential special school tend to score higher W.I.S.C. Performance than Verbal I.Qs. These writers suggest that the differences may relate to the fact that "many of their children come from broken or unstable homes, which do not provide the best conditions for language development." In this connection, Bernstein's hypotheses (1958 and 1959) offer a possible explanation of the underlying causes.

In view of the recent appearance of Form L-M (Terman and Merrill 1960), an appraisal was made of the relevance of the present study to the new form, which seems likely to supercede Form L eventually. As in Appendix 1 of Kennedy-Fraser (1948), and as in the present study for T-M (L), the sub-tests of the new Form L-M were similarly categorized as either 'Verbal' or 'Non-Verbal' (which includes 'Arithmetical'), on the basis of the kind of test responses elicited.

Kennedy-Fraser (1948) reports for T-M (L) the following percentages of 'Non-Verbal' sub-tests (including Arithmetical), for M.A. groups Yrs. II—V, Yrs. VI—VII, and Yrs. VIII—S.A.3, 59 per cent., 19 per cent. and 26 per cent., respectively. For the same three M.A. groups, on Form L-M, the percentages of 'Non-Verbal' sub-tests were found to be 61 per cent., 16 per cent. and 30 per cent., respectively. Thus, as far as the 'Verbal'—'Non-Verbal' dimension in the composition of the new Form L-M is concerned, there would appear to be very little difference with T-M (L), and, in fact, many of the sub-tests are the same. The conclusions reached in the present study would, therefore, appear to apply in general to Form L-M.

V.—CONCLUSIONS AND RECOMMENDATIONS.

(1) E.S.N. boys between 7 and 16 years score better on the Non-Verbal than on the Verbal sub-tests of T-M (L). This is in line with earlier findings on the relative efficiency of intellectually inferior groups on verbal and performance tests—E.S.N., deprived and delinquent children and adolescents. This may well be related to the sociological findings that

(2) E.S.N. boys in this Kent residential school tend to come from homes where the father is unskilled or semi-skilled, i.e., from the lowest two occupation groups ; and from larger families of three or more children.

(3) On this experimental population, there are tentative indications that between the mean ages of 9⅛ years and 15½ years, using T-M (L) tests, ' verbal efficiency ' improves somewhat, while ' non-verbal efficiency ' deteriorates a little. Despite these findings and despite the fact that Roberts Mellone corrections were used throughout, there was a statistically significant decline in I.Q. as measured by T-M (L).

(4) This finding may be explained partially at least by Clarke's hypothesis, that it is due to " the increased verbal weighting at the higher age-levels " (83 per cent. as opposed to 64 per cent.), " since verbal assets are more seriously impaired in such children than are performance assets." This is not to deny, however, that other aetiological factors may be operative in causing such a deterioration. Although an attempt has been made by use of the R.M. correction table to correct for the grosser errors of standardization, yet the factor stressed by Stott (1960), " real changes in the rate of mental development," due possibly to the cumulative effects of a culturally deprivatory environment, or to lowered motivation, or to a lowered level of " mental organization and conceptualization," may also be important.

(5) Whatever the causes, T-M (L) would appear to be a misleading test to use in the assessment of E.S.N. children Although it is a predominantly verbal test, and the proportion of verbal items passed improves slightly between 9⅛ and 15½ years, yet its structure is such that the I.Q. appears to decline. It would appear inappropriate for the dull child from a poor or detrimental environment ; it fails to give an all-round assessment of intellectual capacities ; and the different types of items are inconsistently and unequally distributed through the different mental age levels of the scale.

(6) In view of all the difficulties mentioned in this paper, the test under question would appear to need very careful interpretation when applied to an educationally subnormal population, especially if changes in cognitive level were being considered.

(7) The most efficient type of test for use in the assessment of ' dull ' children would appear to be one where verbal/educational and performance/perceptual items were equally balanced, where separate verbal and performance I.Qs. were obtainable, where all the established ' group factors ' were tapped equally and were measurable quantitatively, and where the different types of items were consistently distributed through all age levels of the scale. (A recent attempt has been made by McCarthy and Kirk (1961) to devise a series of tests which analyze the specific psycho-linguistic abilities and disabilities of subnormal children, so that detailed diagnosis can lead to effective remedial treatment. It is as yet too early to judge the effectiveness of such a differential diagnostic instrument, but the question is considered by Cashdan (1962).)

(8) The problems discussed in this paper might well be elucidated, were it possible for some investigator to do a comparable test-retest study on children at the other extreme of the intelligence range—the ' educationally superior ' population. Valuable confirmatory evidence for the main hypothesis of this paper might also be educed, were some of the earlier studies quoted (Phillips 1940, Thorndike 1940, Scarr 1953, Collman and Newlyn 1958, Hiskey and Sadnavitch 1958, Portwood (private communication) and Stott 1960), subjected to the same sort of statistical analysis.

(9) There would appear to be little difference in the 'Verbal'—'Non-Verbal' composition of T-M (L) (1937), and Form L-M (1960). Thus, the above conclusions and recommendations would equally apply to the new Form L-M, as to Form L.

ACKNOWLEDGEMENTS.—We wish to express our thanks to the County Education Officer, the Principal School Medical Officer and the Children's Officer of the County of Kent and their staffs, for the facilities and information provided; to the Head Master of the School, and the School Medical Officers who did the testing; to Professor A. D. B. Clarke, Dr. N. France, Mr. Gwynne Jones and Mr. P. Portwood for their help and advice. Not least are our thanks due to the 111 children who were the subjects of the study. Opinions expressed in the paper are those of the authors personally and *not* of the Kent Education Committee, or its officers.

VI.—REFERENCES.

ANASTASI, A. (1956). Intelligence and family size. *Psychol. Bull.*, 53, 187-209.

BERNSTEIN, B. (1958). Some sociological determinants of perception. *Brit. J. Sociol.*, 9, 159-174.

BERNSTEIN, B. (1959). A public language. *Brit. J. Sociol.*, 10, 311-326.

BOWLBY, J. (1951). *Maternal Care and Mental Health.* Geneva: W.H.O.

BOWLBY, J., AINSWORTH, M., BOSTON, M., and ROSENBLUTH, D (1956). The effects of mother-child separation—a follow-ups study. *Brit. J. Med Psychol.*, 29, 211-248.

Bull. Brit. Psychol. Soc. (1958). Appendix I—the 1937 version of the Stanford-Binet Scale—a critical appraisal, 35, 13-15.

CASHDAN, A (1962). The intellectual powers of subnormal children. *Educ. Res.*, 4, 84-97.

CLARKE, A. D. B.. and CLARKE, A. M. (Eds.) (1958). *Mental Deficiency—the Changing Outlook.* Methuen.

CLARKE, A. D. B , CLARKE, A. M. and REIMAN, S. (1958). Cognitive and social changes in the feeble-minded—3 further studies *Brit. J. Psychol*, 49, 144-157.

CLARKE, A. D. B., and CLARKE, A. M. (1959). Recovery from the effects of deprivation. *Acta Psychol.*, 16, 137-144.

COLLMAN, R. D. and NEWLYN, D. (1958). Changes in Terman-Merrill I.Qs. of mentally retarded children. *Am. J. Ment. Defic.*, 63

DAVIS, D. R. (1957). *An Introduction to Psycho-pathology.* O.U.P.

EYSENCK, H J. (1953) *Uses and Abuses of Psychology.* Penguin.

GLUECK, S., and GLUECK, E. (1950). *Unravelling Juvenile Delinquency.* New York: Harvard U P.

HISKEY, M. S , and SADNAVITCH, J. M (1958). Minimizing exaggerated changes in Binet ratings of retarded children. *Exceptional Children*, 25, 16-20.

H.M.S.O. (1954). *The 1951 Census Report for the County of Kent.*

HOLLINGWORTH, L. S. (1920). *The Psychology of Sub-Normal Children.* New York.

JONES, H. E. (1954). " The Environment and Mental Development " in Carmichael, L. (Ed.) *Manual of Child Psychology.* New York: Wiley.

KENNEDY-FRASER, D (1948). *The Terman-Merrill Intelligence Scale in Scotland*—U.L.P. Scots. Counc. Res Educ. Pubs. 23

KIRK, S. A. (1948). An evaluation of the study by Bernardine G. Schmidt (1946). *Psychol Bull.*, 45, 321-333.

MACNEMAR, Q (1940). A critical examination of the University of Iowa studies of environmental influences upon the I Q. *Psychol. Bull.* 37, 63-92.

MACNEMAR, Q. (1942). *The Revision of the Stanford-Binet Scale: An Analysis of the Standardization Data* Boston: Houghton-Mifflin.

MCCARTHY, J. J. and KIRK, S. A. (1961). *Illinois Test of Psycho-linguistic Abilities.* Univ Illinois.

MUNDY, L. (1957). Environmental influence on intellectual function as measured by intelligence tests. *Brit. J. Med. Psychol.*, 30, 194-201.

NISBET, J. (1953). Family environment and intelligence. *Eugen Rev.* 45, 31-40.

O'CONNOR, N (1956). The evidence for the permanently disturbing effects of mother-child separation. *Acta Psychologica*, 12, 174.

PARKER, H. T. (1934). *The Deviation of Intelligence in Sub-normal Children.* Melbourne U.P.

Penrose, L. S. (1949). *The Biology of Mental Defect.* Sidgewick and Jackson

Phillips, G. E. (1940). *The Constancy of the I.Q. in Sub-normal Children*—O.U.P. Austr. Conc. Educ. Res. No. 60.

Portwood, P. F. (1958). Progress report of the Sheffield E.S.N. Study. *B. Ps. S. Bulletin* No. 34, 61.

Pringle, M. L. K., and Bossio, V. (1958). A study of deprived children, Part 2—Language Development and Reading Attainments. ***Vita Humana I***, 3/4, 142-170.

Roberts, J. A. F., and Mellone, M. (1952). On the adjustments of Terman-NMerrill Intelligence Quotients to secure comparability at different ages. *Brit. J. Psychol.* (*Stat. Sect*), 5, pp. 65-79.

Sandler, J. (1955). A test of the significance of the difference between the means of correlated measures, based on a simplification of Students *t*. *Brit. J. Psychol.*, 46, 225-226.

Scarr, E. H. (1953). Changes in Terman-Merrill I.Qs., with dull children. *Brit. J. Stat. Psychol.*, 6, 71-76.

Schmidt, B. G. (1946) Changes in personal, social and intellectual behaviour of children originally classified as feeble-minded. *Psychol., Monogr.*, 56, 5.

Sontag, L. W., Baker, C. T., and Nelson, V. L. (1955). Personality as a determinant of performance. *Am. J. Orthopsychiat.*, 25, 555-562.

Sontag, L. W., Baker, C T., and Nelson, V. L. (1958). Mental growth and personality development : a longitudinal study. *Soc. Res. Ch. Def. Pubs.* 23, Antioch, Ohio.

Stott, D H (1960). Observations on retest discrepancy in mentally sub-normal children. *Brit. J. Ed. Psychol.*, 30, 211-219.

Tansley, A. E., and Gulliford, R. (1960). ***The Education of Slow Learning Children.*** Routledge and Kegan Paul.

Terman, L M., and Merrill, M. A. (1937). ***Measuring Intelligence.*** Harrap.

Terman, L. M., and Merrill, M. A. (1960). *Stanford-Binet Intelligence Scale.* Harrap.

Thorndike, R L (1940). Constancy of the I.Q. *Psychol. Bull.*, 37, 167-186.

Vernon, P. E. (Ed.) (1957). *Secondary School Selection.* Methuen.

Wellman, B. L. (1934). Growth in intelligence under differing school environments. *J. Exp. Educ.*, 3, 59-83.

(*Manuscript reeceived 29th, March* 1962.)

THE DEVELOPMENT OF CHILDREN'S IDEAS ABOUT COUNTRY AND NATIONALITY

PART II: NATIONAL SYMBOLS AND THEMES

BY GUSTAV JAHODA
(University of Glasgow)

SUMMARY. The learning of conventional national symbols (national anthem, songs, flags) and the understanding of their significance at various stages were studied. A specially devised picture test was used to explore more subtle aspects of national consciousness. The broad developmental trends and the socio-economic group differences emerging from both parts of the investigation are outlined, and some educational implications suggested.

I.—INTRODUCTION.

IN the literature of the social sciences a good deal of lip-service is paid to the importance of national symbols in the building up of national sentiments. For instance, G. W. Allport (1950) wrote as follows: " As a rule, personal loyalty can adhere to an abstraction only when the abstraction is richly symbolized. Christianity rivets attention upon the cross, nations focus on their respective flags." Few attempts seem to have been made to trace the acquisition of such symbols in young children, apart from the work of Horowitz (1940) and recently Lawson (1963), who concentrated on national flags. The aim of the present study is to initiate the exploration of a somewhat wider range of symbols.

II.—THE NATIONAL ANTHEM.

Six different tunes were played to the children on a tape-recorder in a systematically rotated order. Two of them were Scottish, two Irish, one English, and the National Anthem. They were asked to say of each if they had heard it before; if recognition was claimed, the following questions were asked: Where did you hear it? Do you know what it is? Which country does this tune belong to? Here only results relating to the National Anthem will be given.

Table 1 shows the different kinds of responses. About one-quarter of the youngest children gave no sign of recognition. The second group said they had heard it before, but were either unable to specify the country, or gave absurd answers like " From the country where Herod is " (note the tense), or even " It's from Wagon Train."

TABLE 1
RESPONSES TO HEARING NATIONAL ANTHEM.

	M Schools			*W* Schools		
	6–7	8–9	10–11	6–7	8–9	10–11
Fails to recognize tune	6	—	—	7	1	—
Claims recognition but does not know country, or absurd response	13	4	2	11	6	4
Claims recognition and attributes to Scotland	1	4	2	4	9	9
Ditto and attributes to London or England	3	6	1	—	2	4
Ditto and attributes to Britain	1	10	19	2	6	7
	24	24	24	24	24	24
Per cent. unable to give title of tune	54%	4%	—	87%	12%	—

At this point attention may be drawn to the bottom row of the table, indicating the frequency of ignorance of the name. For this purpose only "God Save the Queen" or "National Anthem" were accepted; of those rejected, five implied awareness that the tune was linked with the monarchy, such as "The royal song" or "The Queen's music." Beyond the age of 7, inability to name the tune was rare, and in general it was found more often among *W* children (P less than ·02).

Turning now to a consideration of errors of attribution, one finds two major trends. The first, most common with *W* children, was to claim the anthem for Scotland. The explanation which suggests itself is that these children are unlikely to have attained a conception of Britain as a whole, and a check of G-stages supports this: 76 per cent. of them were at or below G.III, as compared with only 30 per cent. of those answering correctly (P less than ·001). The causes of the second error, namely mentioning England or London, are less evident. In some cases at least, judging from the replies, it seems to have arisen from an association in the children's minds between the Queen, her residence in London, and 'her music.'

The pattern of development that emerges from the comments (though owing to limitations of time it was not possible to pursue it systematically) may at least be broadly sketched. At earlier ages, the National Anthem tends to be just one particular piece of music among others. The name is largely unknown, but it is connected with such occasions as school ceremonials, Church and Sunday School (e.g., "We get it at concerts"; "In the service in the hall—the big ones know the words"; "It's a hymn"); other sources include the wireless and, especially among *W* children, television.* These trends continued at 8–9, but there were some additions. For *M* children, these consisted mainly of Cubs' or Brownies' displays, with *W* ones 'the pictures' or miscellaneous occasions like "Away up the town when the band was playing." Moreover, the name then being generally known, a definite link with the idea of the Queen had become established. In the case of most of the older (10–11) *M* children, the tune had already taken on the character of a national symbol (e.g., "Well, that can be sung by people all over the British Isles"); with the corresponding *W* group, such progress tended to be much slower, as illustrated by the examples below:

> It was on the T.V., when they stand up and sing. (*Why do they stand up*?) Don't know. (*W*10).
>
> They play it at the start of an international game or something. It could be English, Scottish—anything. (*W*11).

Thus, there is a slow and almost imperceptible transition in the meaning of this music for the children. At first it may be quite unfamiliar or, if recognized, quite wrongly labelled. Later, it comes to be connected first with the person of the Queen, and then with Britain as a whole; at that stage it was almost invariably correctly named. It was striking to observe how older children, who had already grasped this, responded to the questions with a contemptuous laugh, because they felt the answers were so obvious; one would hardly have guessed that their insight is likely to have been a fairly recent acquisition

III.—NATIONAL SONGS

Whilst the National Anthem is confined to more formal or ritual occasions, there are many popular tunes tinged with patriotic sentiments Furthermore, by contrast with the National Anthem, such tunes tend to be associated with the various units which together constitute the United Kingdom. The tracing of development in this sphere was the object of putting the following question : Many people from this country go to live in far-away places and are a long way from home Now imagine they would like to hear three songs : which ones do you think they would choose ?

---

* One-third of the 7-year-old *W* children were able to report that "It's on the T.V. when it's finished."

Answers were divided into three categories. The first consisted of those unable to give the name of any song, together with those who produced irrelevant ones. For all these the question effectively meant " Tell me the names of three songs." The youngest children in this category came out mainly with religious songs (" Gentle Jesus," " Jesus loves me," " Away in a manger," etc.). At a later age, irrelevant songs tended to become increasingly secular, and some divergence between *M* and *W* emerged. " I saw Mum kissing Santa Claus " or " Jailhouse Rock " *versus* " Copelia " or " An adult might like opera "). Also included in this category, because their number was relatively small, were what might be called transitional answers. In these, the child seemed to have some inkling of what was required, but was as yet unable to bridge the gap imaginately by putting himself into the position of someone far from home. Examples of this would be children who suggested such people would wish to hear Italian, Indian or other foreign songs ; one boy said : " We might go to France next year, so I think they should be French ones " ; in Piaget's terms, this would be a typically egocentric statement. All the responses so far are grouped under ' 'Nil or irrelevant " in Table 2.

TABLE 2

SONGS FROM HOME.

	M Schools			*W* Schools		
	6–7	8–9	10–11	6–7	8–9	10–11
Nil or irrelevant response	20	14	5	21	14	4
Understood and produced none or one ..	3	4	7	2	7	7
Understood and produced two or three..	1	6	12	1	3	13

The second category was composed of children who clearly understood what was required (" They will want to hear a Scottish song," or words to that effect), but either could not think of any Scottish song at all, or were able to name only one. There was a small sprinkling of correct non-Scottish answers such as " Carry me home to blighty," which were, of course, credited. The third category was the same, except that the children knew two or three songs.*

One noteworthy feature of this question is the fact that it is one of the few which failed to yield any significant difference between *M* and *W*. From this, it may perhaps be inferred that the establishment of an effective association between native songs and the idea of the homeland is largely independent of differences in both intelligence and social background. It would seem that conceptually the presence of a global, diffuse and undifferentiated notion of, in this case, Scotland is sufficient for the formation of the link. Thus, none of the

---

* Those most frequently mentioned included : Auld Lang Syne, I belong to Glasgow, Scotland the Brave, Roamin' in the gloamin', Loch Lomond, A Gordon for me, Song of the Clyde, A hundred pipers, Ye banks and braes, My heart's in the Highlands.

G.I. children understood the problem ; however, the superiority of G.III over G.II was slight and not significant (P between ·30 and ·20). It is also likely that children in both kinds of environments are about equally exposed to patriotic songs, learning them at roughly the same rate.

IV.—THE UNION JACK.

Nine coloured cards picturing, among others, the Union Jack and the Stars and Stripes, were spread out in front of the children. At the outset they were told to point out those they had never seen before, and these were removed. Next, they were asked the names of the remainder, and the countries to which they belonged. Finally, children unable to comply were informed that some of the flags were Scottish, with the request that they should try and pick them out. The last move turned out to have been a mistake, hence, unfortunately, no useful material was obtained regarding the Lion Rampant and St. Andrew's Cross.* The report on the findings will, therefore, be centred on the Union Jack.

Results are set out in Table 3. First of all, there were two 6-year-old *W* children who placed the Union Jack among those they had never seen : they are included in the top row. It will be noted that the children were generally able to name the Union Jack correctly before attributing it to Britain as a whole rather than claiming it for Scotland only. The *M* children were significantly more successful on both counts (P less than ·01).

TABLE 3

THE UNION JACK.

	M Schools			*W* Schools		
	6–7	8–9	10–11	6–7	8–9	10–11
Does not know country or mistakes other than Scotland	5	1	1	15	5	2
Calls it Scottish	14	8	4	4	11	9
Correct	5	15	19	5	8	13
	24	24	24	24	24	24
Per cent. knowing the name of flag	42%	92%	96%	29%	54%	83%
Number of children able to identify U.S.A. flag	3	14	22	4	11	15

The youngest children were apt to link the flag with concrete situations in their spontaneous comments, as shown below :

> It's the flag you always put up when the Queen comes. (*M*6).
>
> That's the one you get at the circus—my daddy buys it for me. (*What country*?) It's the Glasgow flag. (*W*6).

The Union Jack was often called " the Queen's flag," indicating an association already seen in connection with the National Anthem. Beyond the age of 7, grosser mistakes were mainly confined to *W* children ; apart from calling a variety of different flags " Union Jack," they sometimes used the expression in a completely uncomprehending manner :

---

* This was because of a pervasive tendency to call the whole of the remainder " Scottish," whereby they were bound to score by chance alone.

> I think that's the Jack flag. (*What country* ?) Don't know. (*W*9).
> Yes, that's the Union Jack. (*What country* ?) Is it France ? (*W*9)

In contrast, to this, the bottom row of Table 3 shows that *W* children did not do significantly worse than *M* ones in attributing the Star and Stripes to the U.S.A. It might be added that a good many more recognized the flag without being able to specify the country. Here is a revealing instance :

> That's General Custard's flag ; I've seen him on T.V. ; he fights the Indians. (*W*7).

Lastly, a few impressionistic comments will be offered regarding the Scottish emblems. They appeared familiar to many children, though few even among the older children were able to name them correctly. The pervasive influence of television again seems to have made itself felt, as several children mentioned it in connection with the Lion Rampant, this being the symbol of Scottish Television. Some of the *M* children who did best explained that they had learnt it with the Cubs or Brownies. On the whole, there was little evidence that, within the age-range studied, the Scottish emblems functioned as effective symbols of Scottish nationality.

V.—RECOGNITION OF SCOTTISH PICTURES.

This procedure was devised with the object of assessing some of the subtler aspects of children's knowledge of their country. They were presented with sets of three pictures, reproduced in Plates I–V, and required to indicate for each set which—if any—of the pictures is connected with Scotland.

PLATE I—COSTUMES

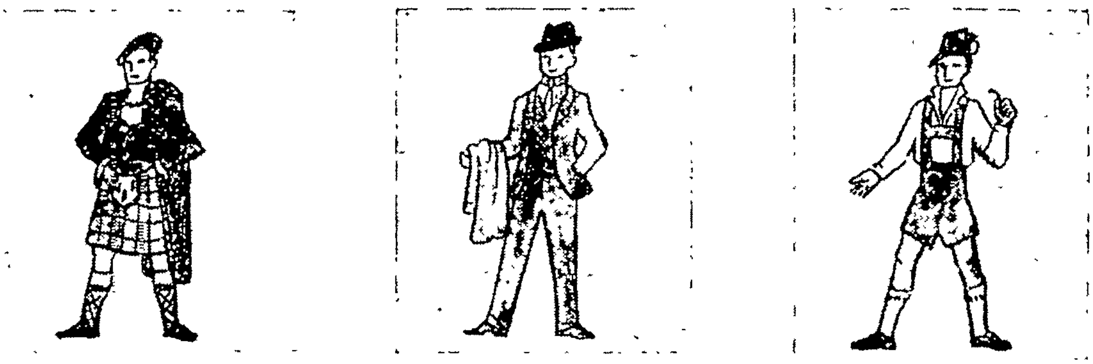
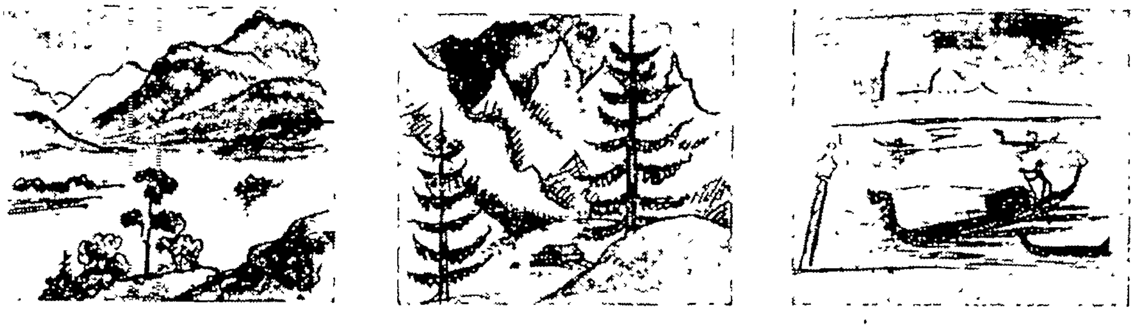


PLATE II—LANDSCAPES



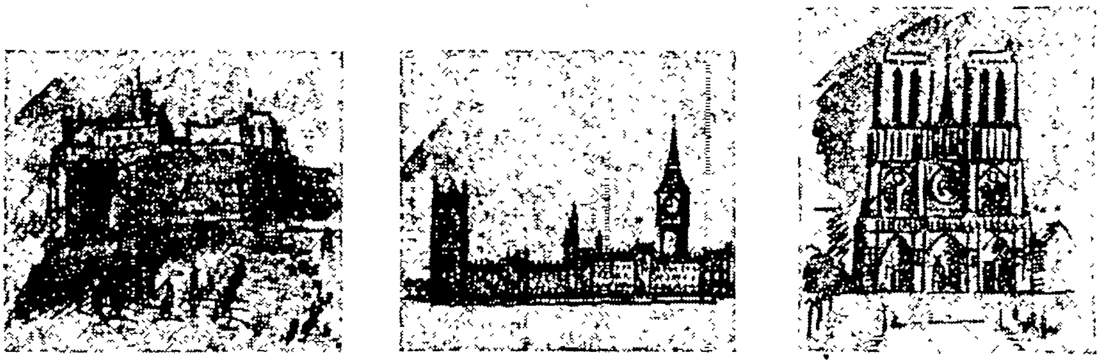


PLATE III—BUILDINGS

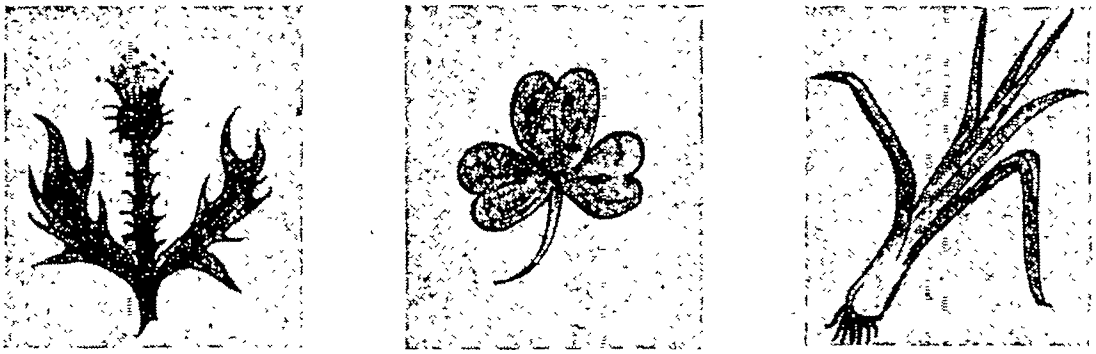


PLATE IV—EMBLEMS

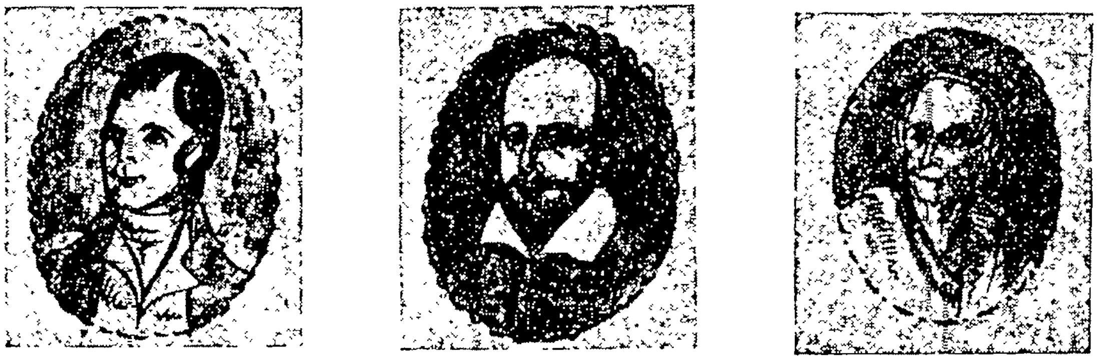


PLATE V—PORTRAITS

Table 4 gives the results in a somewhat different form, being confined to the number of successes for each sub-group of children. No correction was made for possible chance successes, as it was evident from the children's comments that their number was negligible. The first set, mainly intended as an encouragement, dealt with dress and costumes. The great majority of children had no trouble with this. With the landscape set, where a typical Scottish scene had to be picked out, *W* children were conspicuously less successful (P less than ·001). The difference probably reflects differential opportunities for travelling around the countryside, and possibly also differences in the extent to which landscapes decorate their homes, though the latter is entirely speculative. Similar considerations would apply to the next set, three buildings including Edinburgh Castle (P less than ·02). It is noteworthy that several *M* children were able to identify Notre Dame in Paris, one of the ' distractors.'

After this came a set of conventional emblems, and here the rates of success in singling out the ubiquitous Scottish thistle differed even less, though still significantly (P less than ·05). The last set, featuring " Rabbie " Burns in company with Shakespeare and Voltaire, produced a surprise. There was no significant *M–W* difference, either in correctly choosing Burns as the Scottish one, or even in knowing his name. Evidently Burns constitutes one of the most potent Scottish national symbols, as early as childhood. Below are some revealing answers to the question " How do you know this is Burns ? " :

> My grannie's got a picture of him. (*M*8).
> Because we have a calendar with his face on it. (*M*7).
>
> When your mummie's in hospital you go into a fish-and-chip shop for your dinner ; and when I looked up I saw a picture of him looking at me. (*W*6). (N.B.—This boy recognized neither the National Anthem nor the Union Jack.)
>
> On the T.V. news—we saw his cottage. (*W*7).

It would seem that Burns' presence in Scottish life cuts across the barriers of social class, and that is no doubt how he would have wished it to be.

TABLE 4

RECOGNITION OF SCOTTISH PICTURES.

(Numbers of children out of 24 who succeeded on each task.)

	M Schools			*W* Schools		
	6–7	8–9	10–11	6–7	8–9	10–11
Costumes	18	23	24	22	23	23
Landscapes	20	21	23	8	16	18
Buildings	15	21	21	9	15	19
Emblems	18	19	22	11	18	19
Portraits	14	16	22	10	20	22
Percentage of children able to name Burns	29%	46%	88%	21%	58%	88%

VI.—ATTITUDES TO OTHER COUNTRIES AND PEOPLES.

The findings relating to this are reported in detail elsewhere (Jahoda, 1962), and this section, therefore, offers merely a sketch of broad trends and some features of special interest in this country.

The children's growing awareness of more distant nations began with casual glimpses of unusual picturesque detail divorced from any general background. This was sometimes combined with concrete and homely, though equally fragmentary information gained through such personal channels as relatives abroad. The subsequent process of differentiation and organization, too complex for brief summary, was studied in some detail in relation to America and Russia.

Attitudes tended to follow a distinct pattern. Those of the youngest and/or least mature children were determined by some arbitrary and usually concrete detail that happened to have created an impression (e.g., " Don't like New Zealand because the butter is too salt "). Actually, these are probably too evanescent to deserve the term 'attitudes,' and are mentioned only because they are characteristic of an early mode of thought. Apart from the attraction of 'holiday countries,' the more stable sources of favourable attitudes seemed to be again the existence of personal ties, or a belief in the similarity of the people to oneself, resulting in a feeling of affinity ; this last statement, it should be added, is somewhat speculative and could not be directly verified from the data. Unfavourable attitudes, from a surprisingly early age, have their roots in past wars, but among the oldest or most mature children these are overshadowed by present political conflicts.

On the salient, and somewhat disturbing findings, was the ethnocentricity exhibited by some *W* children. This took the form of colour prejudice, and a marked tendency to look down on people in other countries. It is interesting to note, however, that their attitude to the English appeared to be influenced by considerations of social class. Thus, a higher proportion of *W* as compared with *M* children thought the English 'nicer' in behaviour (e.g., " In London the men are not so rowdy when they're drunk ") ; moreover, they tended to identify an English accent with middle class speech, and to regard the Scottish tongue as a working class one, as may be illustrated by two examples :

> Well, the English say " Good jolly, this is a jolly fine day." The Scottish say " Aye, it's a grand day." The English are a lot of toffs. (*W*9).
>
> They talk different—kind of polite, kind of Cockney . . . (*W*11).

Lastly, a few general remarks on attitudes towards the English. From the way the children talked about them (and it must be remembered that will largely reflect the expressed opinion of their elders) it would seem that the traditional antagonism is very much on the wane. There were hardly any disparaging remarks about the " Sassenachs "*, and when questioned about dislikes, not a single one mentioned England or the English.

VII.—CONCLUSIONS.

The outstanding impression one gains from surveying the children's development, in line with the conclusions of investigators concerned with other areas of children's thinking, is that of the great intellectual distance traversed within the span of a few years. At the outset, their responses reveal an almost complete ignorance about the wider geographical and social world surrounding them; some five years later the outlook of most of them is no longer fundamentally different from that of mature adults. During the intervening years, they have acquired a set of conceptual tools enabling them to organize their environment meaningfully and to make the main distinction conventional in our society.

---

* Asked to give another name for the English, only three *M* children were able to produce this ancient word.

The details of this process have been described in this present study. However, it was necessary to treat the various aspects more or less in isolation, leaving aside the question as to whether there may not be some common thread running through the advances in seemingly divergent spheres. An analysis carried out with the material as a whole indicates that the central element is probably to be found in the extent to which the child has succeeded in differentiating and ordering ever-widening sections of his geographical environment; and this can be at least roughly assessed in terms of the G-stages. It would appear that the particular G-stage reached is an index of the limit to the amount of progress possible elsewhere; children often lag behind, failing to realize their potential achievement, but there is no indication that they can exceed it in genuine understanding, as contrasted with mere parrot recital.

Among the youngest children investigated, nearly nine-tenths had not advanced beyond G.II. It is thus legitimate to generalize about the age-group 6–7, because their outlook was fairly uniform. After the age of 7, the increasing diversity of G-stages precludes any general statement, which would be misleading.

It may be recalled that G.I. denotes a perspective confined to the child's immediate surroundings—home, school and play streets in which they spend their everyday life. With G.II an extension has taken place, involving some apprehension of Glasgow as a unitary whole, to which the child refers himself and of which he feels himself to be a part. As yet, however, no conception has been formed about what lies outside. From this starting-point, an inventory of the ideas characterising a majority of the children aged 6–7 will be made.

First of all, the meaning of even the simplest geographical words tends to be extremely vague, or altogether mistaken. Thus ' town ' is liable to be applied to any aggregate from streets upwards. This reflects not only inadequate command of this particular concept, but at the same time also inability to classify known geographical names in any coherent order; both are really two sides of the same coin. Asked " what is your country? " the dominant answer was " Glasgow "; one replied " I don't live in a country, I live in a close.* Whilst they have heard the *word* " Scotland," and if G.II usually realize that it is a place of some kind, they cannot locate it properly in space; it is thought of as being " away " somewhere. Requests to say something about Britain elicited a collection of absurdities. The most they achieved, with three exceptions, was to put forward notions implying recognition that Britain is some kind of geographical entity, though invariably a remote one. Consequently, symbols relating to Britain as a whole, such as the National Anthem or the Union Jack, were at best attributed to Scotland; a majority of children in G.I failed to recognize them altogether.

When asked whether they were British, a majority of the children in G.I and G.II replied negatively; the same answer was given by nearly all children in G.I to the question as to whether they were Scottish; but two-thirds of those in G.II affirmed that they were Scottish. In view of the evident confusion prevalent even among the latter, this result may appear surprising: why should they be ready to acknowledge being Scottish, when they clearly fall short of understanding the relationship between Glasgow and Scotland? In order to explain this, one has to abandon the exclusive concern with purely cognitive aspects. These children are growing up in a symbolic environment (i.e., that part of their environment consisting of words, signs, etc.) where Scotland and things Scottish are bound to figure rather prominently. Hence, even before they are capable of building up a coherent conceptual framework serving to categorize their geographical surroundings, they will become familiar with such words as " Scotland " and " Scottish," as well as to a lesser extent " Britain " and " British." Moreover, some broad and loose associations, normally with a favourable emotional tinge, will tend to be established between these words and certain items of their experience. This was often apparent in earlier pages: children were reminded by the word " Scotland " of their family, of Highland Dancers, and so forth. An accumulation of such associations leads to a readiness to acknowledge themselves as " Scottish " ; but until such time as an adequate concept of " Scotland " and " Scottish " has crystalized, such quasi-identifications remain unstable and apt to be withdrawn when challenged.

---

* Close: strictly an entrance or passage, but extends to the common stairs of a tenement, or the building as a whole.

The purpose of labouring this point at some length is to show again how young children are sometimes able to use these words in a formally correct manner, when more careful probing would reveal that conceptual grasp still eludes them. In fact, during the first two or three years of their school career, most children have hardly begun to make sense of the wider geographical, social and political setting around them. It will not have escaped the readers' attention that it was necessary to qualify statements about the youngest children, because even they were far from uniform. Thus, one 6-year-old was already in G.IV, whilst at the other end of the scale, two 10-year-olds had not moved beyond G.II. This variability of individuals cannot be stressed too strongly, because it is in danger of being overlooked when one thinks in terms of age trends, and this applies most of all to the intermediate ages. A majority of the youngest children are likely to be in G.I or G.II, and most of the oldest in G.III or G.IV, but those aged 8–9 are distributed over the whole range. This holds true for children from *M* and *W* schools taken separately, as well as for the total sample.

In general, the children from *M* schools were found to be significantly in advance of those from *W* schools; however, when *M* and *W* children were matched for intelligence test score, the difference disappeared. It was, therefore, concluded that this factor accounts for a large part of both individual variations and group differences.

VIII.—SOME EDUCATIONAL IMPLICATIONS.

" He that hath names without ideas, wants meaning in his words, and speaks only empty sounds." It would be entirely in the spirit of John Locke to add that the passage from names as more or less empty sounds to names backed by fully elaborated concepts is the outcome of a gradual learning process. However, it is not always easy to know whether or not a child who has a name speaks more than empty sounds. An adult may readily be led to assume that notions self-evident to him will also be within the reach of the child. One might again turn to Locke for an example. : " The relation also that things have to one another in their *places* and distances is very obvious to observe ; as above, below, a mile distant from Charing Cross, in England, in London." An analysis of the concepts underlying such statements, along the lines sketched in the present study, will indicate that to a child such relationships may be far from obvious.

What bearing does this have on the teaching situation ? The simple answer would be that the teacher should not jump too far ahead of the children's capacity for adequate comprehension ; but this is hardly a very helpful recommendation. In practice, the teacher could not hope to ascertain the stage of conceptual development reached by each child, and even if this could be done, some compromise would be unavoidable. More effective action would be a reconsideration of the manner of teaching such subjects as Geography and History. An examination of the Schemes of Work for teaching these in primary schools, keeping in mind the findings reported, will reveal the presence of numerous assumptions that are likely to be invalid ; for instance, what kinds of ' simple geographical terms ' are regarded as having to be taught, and which ones seem to be tacitly taken for granted ?

Another possible suggestion is that some account be taken in planning the curriculum of what appears to the dominant, if not universal sequence of concept development from the immediate vicinity outwards. This would rule out, say, a leap from plans of the school room and adjacent streets to the globe or map of the world. The main aim would be a consolidation of the series widening out from the home town to Britain as a whole.

Finally, it is important to guard against a possible misconception, arising from the general discussion of developmental stages. It has been stated that a child must attain a certain level of intellectual maturity before he is ready to move from one stage to the next ; he must be capable of performing the appropriate logical operations as well as of handling spatial relationships. Thus, subject to the wide range of individual variation already emphasized, there is probably a ceiling to what a given child can understand at a particular time. However, it might be wrongly inferred from this that there is relatively little the teacher can do to help a child along. Such a view presupposes that most children more or less automatically advance to the limits of their potentialities, and this is, unfortunately, very far from the truth. Even in the course of the actual investigation, as an unintended consequence of merely trying to answer questions of a kind with which they were not normally faced, some children had a flash of insight ; things suddenly fell into place, and they reached a new level of understanding. If this can be brought about so casually and incidentally, it is likely that teachers could achieve far more by a systematic effort. It is hoped that the present investigation may make some limited contribution by pointing the way.

ACKNOWLEDGEMENTS.—Thanks are due to Dr. H. Stewart Mackintosh, Director of Education, and to Heads of Schools for their help with the research, and to Mrs. Diana Buchanan who prepared the material for the picture test. Most of all, I am indebted to my colleague Alistair Chalmers who carried out the main fieldwork and helped with the analysis. The research was supported by a grant from the Court of Glasgow University.

IX.—REFERENCES.

ALLPORT, G. W. (1950). In T. H. Pear (Ed) : *Psychological Factors in Peace and War.* London : Hutchison, p. 153.

HOROWITZ, L. (1940). Some aspects of the development of patriotism in children. *Sociometry*, 3, 329-341.

JAHODA, G. (1962). Development of Scottish children's ideas and attitudes about other countries. *J Soc. Psychol.*, 58, 91-108.

LAWSON, E. D. (1963). Development of patriotism in children—a second look. *J. Psychol*, 55, 279-286.

(*Manuscript received* 20*th August*, 1962.)

THE SELF CONCEPTS OF TRAINEE-TEACHERS IN TWO SUB-CULTURES

BY A. S. PHILLIPS
(*Department of Education, University of the West Indies*)

SUMMARY. (1) The aim of this study was to compare the Self Concepts of training college students from two different sub-cultures, viz., England and Jamaica

(2) Unstructured compositions by students from both countries, on the topic—" I myself : the person I am ; the person others think I am ; the person I would like to be," were analysed to reveal the categories students use in describing themselves. Some interesting qualitative differences appeared though students from both countries used the same categories in their self description.

(3) Objective tests for measuring the Self Concept of students were constructed and administered to other students in both countries. The following comparisons between the two groups emerged :

(*a*) Jamaican students, both male and female, of the regular age range as well as the mature, rated themselves significantly higher on the Cognized and the Ideal Selves, and were more Self Accepting than their English counterparts ;

(*b*) English male students rated themselves at approximately the same level as English female students. But, in the Jamaican samples, the female students tended to rate themselves higher than the men, and to be less Self Accepting.

(*c*) There were no singificant differences between the Self Concept scores of ' regular ' English and ' mature ' English students. But the ' mature ' Jamaican students cognized themselves significantly lower than the ' regular ' Jamaican students

I.—INTRODUCTION.

WITHIN recent years the Self Concept of the individual has been receiving more and more attention from psychologists. It has come to be regarded as the vital part of personality structure, exercising a profound effect on behaviour and on happiness. According to Lecky[4] one of the functions of the Self is to produce an harmonious integrated personality. This is achieved largely by an acceptance of oneself as one is, not in a sense of complacency, but with a full knowledge of the things about oneself one cannot change, and of the endowments one has to work on. Such understanding leads to mental health. Besides, our general behaviour is to a large extent, regulated and organized by what we perceive ourselves to be.

If the individual, therefore, is to lead a reasonably happy, normal, well-adjusted life, free from fears and tensions, from feelings of constant self-reproach and utter inadequacies, it is fairly clear that he should acquire as early as possible, a Self picture that is adequate and realistic, and which will enable him to think of himself in a satisfactory manner. This obviously is a task for the agencies of education. Staines[5] has shown that satisfactory changes can be made in the Self Concepts of children, without any lowering of achievement in the ordinary school subjects, by merely adapting the methods of instruction If the Self Concepts of children are of such vital importance, then the Self Concepts of trainee teachers must be of equal importance and worthy of further study. This research was, therefore, designed to probe the Self Concepts of teachers in training. This was done in two different sub-cultures, viz., Jamaica and England, and comparisons were made between the two groups.

II.—THE JAMAICAN SOCIETY.

The Self Concepts of students from the Jamaican and English sub-cultures might be expected to differ widely as the social forces in these two countries have been so varied. Jamaica has been an English colony for over 300 years. Many races compose the population, but the two dominant ones are the white and the negro. From the beginning of its colonial history, these two races were brought together with very clearly defined places in the social hierarchy. The white people owned all the property and all the wealth and were thus in a position of authority and prestige. The negroes were, in the beginning, merely slaves without rights—the chattel of their masters. Emancipation when it came, secured the physical freedom of the slaves, but the pattern continued. The coloured people were made to feel inferior in so many ways, some subtle, some more brutal, that it was virtually impossible for them to doubt their status. This was strengthened by the processes of education, when education did arrive for the coloured children. Its pattern was copied almost entirely from the United Kingdom. The syllabus of the schools was that designed for the United Kingdom. The standards and the ideals were theirs. Even the illustrations and the pictures presented to children were those prepared for children in England. Only in rare circumstances would a child see a negroid type depicted in an illustration, and even then it was usually conforming to the stereotype of ' backward native.' Even the language they were taught was full of expressions equating black with bad or undesirable, and white with good or desirable. The image held up to the child then, particularly in relation to physical characteristics, would be that of the European, and the European type would certainly form the Ideal Self picture.

This probably explains the coloured person's rejection of the negroid part of his ancestry (there has been much intermixture between the races in Jamaica) and his desire to escape from it, as typified by such practices as straightening the hair and using bleaching creams to lighten the complexion.

Though the more obvious forms of discrimination no longer obtain in Jamaica and there is fairly general equality of opportunity for all races, nevertheless, the attitudes still persist, and might be expected to produce Self Concepts full of unworthiness and inadequacy and marked by excessive rejection of Self. There should, therefore, be marked contrasts with the Self Concepts of training college students from England, where no such forces appear to operate.

III.—SOME UNSTRUCTURED COMPOSITIONS.

As a first step, groups of training college students in England and Jamaica were asked to write unstructured compositions on the topic : " I myself : the person I am ; the person others think I am, and the person I would like to be." These three sub-divisions of the main topic were taken as corresponding to the three levels of the Self, viz., the Cognized Self, the Other Self, and the Ideal Self. The aim was to discover the categories that students use in reporting on the Self Concept at these three levels, and also to provide qualitative evidence as to the differences in the way students from both countries cognize themselves.

There was evidence in the essays that students were quite honest in giving their estimates of themselves, perhaps because they had been assured of anonymity. Remarks such as the following were quite frequent :

> " If this were to be read by any one in this college it would be quite impossible for me to write this."

or

> " I find it difficult for me to describe myself, but I shall try to be perfectly honest."

There was some evidence too, that the exercise tended to make the students somewhat uncomfortable, e.g.:

> " I am, at the moment, not liking putting down all these thoughts about myself, because it injures my pride to see myself in such a bad light "

or

> " This hurts, but this is the truth."

An analysis of the essays showed that students in both sub-cultures used more or less the same categories in defining themselves. These categories, with the percentages from each group that gave replies in each category (for the level of the Cognized Self) are set out in Table 1.*

TABLE 1

PERCENTAGE OF STUDENTS WHO MENTIONED ITEMS IN VARIOUS CATEGORIES AT THE LEVEL OF THE COGNIZED SELF.

Categories	English				Jamaican			
	Regular		Mature		Regular		Mature	
	M	F	M	F	M	F	M	F
N—	60	60	—	27	69	25	—	100
1. Physical Characteristics	35	26·6	—	14·8	21·7	32	—	55
2. Clothes, Dress and Grooming	5	13·3	—	11·1	7·3	8	—	26
3. Material Possessions	11·6	13·3	—	7·4	5·8	4	—	13
4. Home and Family	11·6	43·4	—	25·9	18·8	16	—	60
5. Animals and Pets	0	10	—	7·4	0	0	—	1
6. Intelligence	43·3	20	—	7·4	5·8	8	—	12
7. Personality and Character	76·6	93·3	—	100	84·1	92	—	93
8. Philosophical Category	40	38·3	—	25·9	71	48	—	22
9. Performance Category	73·3	73·3	—	39·3	56·6	64	—	86
10. Social Relationships	75	91·6	—	81·5	62·3	68	—	92
11. Status and Role	8·3	21·6	—	3·7	4·4	4	—	8
12. Wants, Goals and Aspirations	6·6	20	—	7·4	16	8	—	16
13. Aesthetic Life	18·3	31·6	—	25·9	10·2	16	—	24
14. Miscellaneous	0	5	—	0	2·9	4	—	2

Certain qualitative differences in the way in which students in each group described themselves also emerged. For instance, there was a tendency for the English students, especially the male students, to denigrate themselves. For example, one man wrote:

> " I make a hell of a lot of noise; I talk too much, drink too much, lie too much, but I can't work up a conscience about any of these things."

On the other hand, the Jamaican group tended to be largely uncritical of themselves, except in minor facets of personality. They tended to express themselves in superlatives, and to see themselves through very rosy spectacles, e.g.:

> " (I am) a warm-hearted person (and) have some of the finest qualities of which human beings are capable."

---

* For further details, see Phillips, A. S—" A Study of the Self Concepts of Selected Groups of Training College Students in their relation to other variables in the Teacher Training Situation." Unpublished Ph.D. Thesis, London University.

> " I am . . . person of more than average ability. I am straightforward in all my dealings, thus causing my character to be beyond reproach."

Certain similarities were evident among the various groups, however. Table 1 shows the high percentage in all groups that gave responses in the categories of ' Personality and Character,' ' Performance,' and ' Social Relationships.' This is in accordance with Jersild's[2] claim that mature people tend to see themselves in terms of inner forces rather than of external qualities. The generally low frequency of responses in such categories as ' Material Possessions,' ' Clothes, Dress and Grooming,' ' Status and Role,' etc., applied to all groups in both subcultures.

Much greater emphasis was placed by both groups on the Cognized Self, rather than on the other two levels of the Self.* This may have resulted merely from the fact that the Cognized Self appeared first in the essays and so received far fuller treatment, though many students admitted that they had great difficulty in knowing what other people thought of them. Also, some, particularly among the Jamaicans, claimed that they would not wish to be any different from what they were.

IV.—CONSTRUCTION OF AN OBJECTIVE TEST FOR MEASURING THE SELF CONCEPT.

In order to secure more objective and quantifiable measures of the Self Concept, a test of the self rating type was constructed. This was divided into two parts measuring respectively the Cognized Self and the Ideal Self. It was decided to investigate only those categories at these two levels of the Self, as were actually reported on by the students in their essays. Twelve categories only were investigated, as one, viz., ' Animals and Pets ' did not appear to be of sufficient universality to warrant inclusion. Later, as a result of statistical analysis, another category, that of ' Aesthetic Life ' was also omitted, leaving eleven categories in the final form of the test.

The individual items in the test consisted of statements made by the students in their essays. The main criteria for selection of these statements were that they had to be applicable to people of different races, living in different countries, to both sexes, and to people of varying ages. Also, items had to be positively framed, that is, indicating possession of a desirable quality. Students were asked to rate themselves on a five-point scale (from 0–4) as far as each item was applicable to them. This gave a score for each item, and as items were grouped into categories, a score for each category. A summation of all items gave a total Self Concept score for the level of the Cognized Self. For the level of the Ideal Self, students were asked to rate themselves on these same items, not as they thought they were, but as they would like to be. The algebraic difference between the scores on the Cognized and the Ideal Selves (the discrepancy score) gave a measure of Self Acceptance.

An unequal number of items was selected for each category as some categories were regarded as of wider conception than others. After the first trial run, it was apparent that certain categories hung together, covering more or less the same area of the Self Concept. It was, therefore, decided to group categories into areas. Evidence for such grouping was provided by correlating scores for various categories with each other. Next, the items forming an area were tested for evidence of scalability according to the Guttman technique, and this revealed that they did, in fact, form a scale. On this basis, five clear areas of the Self Concept were revealed, viz. :

(1) Possessions and Desires.

---

* See Phillips, A. S., *loc. cit.*

(2) Personality, Character and Beliefs.
(3) Social Aspect.
(4) Performance Aspect.
(5) External Appearance.

All further comparisons were done in these areas.

The reliability of the test was determined by a re-test after an interval of three weeks, when a product-moment correlation coefficient was calculated. This gave an overall figure of +·87 for the Cognized Self and +·80 for the Ideal Self. This was regarded as adequate.*

V.—COMPARISON OF STUDENTS FROM ENGLISH AND JAMAICAN COLLEGES.

Table 2 gives the numbers and average age of students in each sample. 'Regular' students were taken as those between 18 and 24 years; 'Mature' students were taken to mean those over 25 years. Unfortunately, it was not possible to obtain large enough samples of mature male students in either country.

TABLE 2

NUMBERS AND AVERAGE AGE OF SAMPLES.

Samples	Numbers	Average Age
Regular English Female	58	19 years 4 months
Regular Jamaican Female	61	20 years 4 months
Regular English Male	51	20 years 8 months
Regular Jamaican Male	47	20 years 11 months
Mature English Female	11	36 years 4 months
Mature Jamaican Female	98	30 years 5 months

In comparing scores for training college students from both sub-cultures, only total self concept scores are given here.† These are set out in Table 3.

TABLE 3

MEAN SELF CONCEPT SCORES.

	Jamaican			English				
	N	M	σ	N	M	σ	C R.	Sig
COGNIZED SELF								
Regular Female	57	176·50	12·80	58	137·30	19·30	12·90	·001
Regular Male	43	172·58	10·96	49	143·26	18·70	9·25	·001
Mature Female	88	166·32	17·00	11	143·82	20·80	4·05	·001
IDEAL SELF								
Regular Female	49	195·26	9·45	58	172·87	12·90	10·02	·001
Regular Male	40	181·75	14·62	48	172·63	17·80	2·63	·01
Mature Female	84	190·61	10·70	11	172·45	15·60	4·70	·001
SELF ACCEPTANCE								
Regular Female	46	17·36	10·91	58	35·02	14·48	7·13	·001
Regular Male	38	10·16	10·10	48	29·92	14·50	7·40	·001
Mature Female	76	24·25	13·70	11	27·91	13·50	·61	N S.

---

* For complete test, see Phillips, A. S., *loc. cit.*

† For full comparison, see Phillips, A. S., *loc. cit.*

It had been expected that the Jamaican students would cognize themselves less adequately, that is lower than, their English counterparts, and be less self accepting (i.e., have a higher discrepancy score). It was revealed that the Jamaican students, both male and female, those of the regular age range, as well as the mature female students, rated themselves significantly higher on the Cognized as well as on the Ideal Self. They also had a lower discrepancy score (i.e., they were more self-accepting) though this difference does not reach significance for the mature female group. These trends are generally true for all areas of the selves.

It is not easy to explain the higher rating achieved by the Jamaican student over his English counterpart. Perhaps it has its roots in the awakening national and racial consciousness evident in Jamaica over the last twenty years. During this period has come a more advanced political constitution for the island, and a widening of opportunities for its dark skinned inhabitants. An unofficial colour bar, which operated prior to 1940, has disappeared. Generally speaking, the negro has acquired a new pride in himself. But this does not explain why the Jamaican students, even the much older ones, should cognize themselves so much higher than (instead of more nearly equal to) the English students. Another possible explanation is that defence mechanisms are at work. Many forces have been trying to impress on the Jamaican, his unworthiness and inadequacy. But he cannot live with a self picture as unlovely as this. So he seeks reassurance by conceiving of himself as much better than he actually is. Hilgard[1] commenting on the operation of the mechanism in the Self Concept, points out that the need for self deception arises because of a more fundamental need to maintain or restore self esteem.

Perhaps, however, the explanation lies simply in the differing experiences of the respective groups. The Jamaican samples were drawn from a level of the population which, by and large, did not have many opportunities for getting ahead and achieving success. But these particular ones had achieved greater success than their peers. Thus, they saw themselves as more intelligent, of better personality, able to achieve more, etc. On the other hand, their English counterparts, were less successful than their peers from grammar schools who had gone on to university or entered other more highly esteemed professions. It should be noted, too, that though the Jamaican samples had a higher Ideal Self, they were, in fact, more Self Acceptant.

TABLE 4

SEX DIFFERENCES IN THE SELF CONCEPT.

	Men			Women				
	N	M	σ	N	M	σ	C.R.	Sig.
COGNIZED SELF								
Jamaican—Regular	43	172·58	10·96	57	176·50	12·80	1·64	N.S.
English—Regular	49	143·26	18·70	58	137·30	19·30	1·49	N.S.
IDEAL SELF								
Jamaican—Regular	40	181·75	14·62	49	195·26	9·45	5·03	·001
English—Regular	48	172·63	17·80	58	172·87	12·90	·078	N.S.
SELF ACCEPTANCE								
Jamaican—Regular	38	10·16	10·10	46	17·36	10·91	3·13	·01
English—Regular	48	29·92	14·50	58	35·02	14·48	1·81	N S

It had been expected that there would be no difference in the scores of men and women on the Self Concept scales at either levels of the Self or in Self Acceptance in either sub-culture. It was found that this was true, for the English samples, but not for the Jamaican samples. For the Jamaican samples, though the difference in the total scores of men and women did not reach significance, women cognized themselves significantly higher than men on three of the five areas making up this total score. On the Ideal Self women were significantly higher than men, and they were less Self Acceptant.

It is possible that the explanation of the difference in Self Concept of Jamaican men and women lies in the fact that the Jamaican society is somewhat matriarchal in structure, though in certain respects it is patriarchal of the Victorian flavour.* Women very often have to assume the role of both mother and father in the family They are very often the mainstay and general provider. It is not surprising, therefore, that they should conceive of themselves in higher terms than their male counterparts. But perhaps, they are not satisfied with playing this role. This might account for their being less Self Acceptant than men. No similar pressures appear to be operating in the English population.

TABLE 5

AGE AND THE SELF CONCEPT.

	Regular			Mature				
	N	M	σ	N	M	σ	C R.	Sig.
COGNIZED SELF								
Jamaican Female	57	176·50	12·80	88	166·32	17·00	4·17	·001
English Female	58	137·30	19·30	11	143·82	20·80	1·02	N S.
IDEAL SELF								
Jamaican Female	49	195·26	9·45	84	190·61	10·70	2·60	·01
English Female	58	172·87	12·90	11	172·45	15·60	·10	N.S
SELF ACCEPTANCE								
Jamaican Female	46	17·36	10·91	76	24·25	13·70	3·08	·01
English Female	58	35·02	14·48	11	27·91	13·50	1·51	N.S.

It was believed that by the time people have reached adult status, the Self Concept would have become relatively stable and would consequently change very little in succeeding years. It was, therefore, hypothesised that there would be no significant differences in the Self Concepts and Self Acceptance of 'regular' and 'mature' training college students in either culture. It was found that this was true of the English samples but not of the Jamaican samples. There are some interesting differences in the direction of change in the different areas however.†

It seems reasonable to conclude that by early adulthood the personality structure of students in the English sub-culture has achieved greater stability than the personality structure of students in the Jamaican sub-culture. Consequently, while the Self Concepts of English students change little with succeeding years, Jamaican students are still changing and modifying their concepts of themselves. Perhaps they are less mature emotionally than their English counterparts.

* See Kerr, M.[3]

†For full exposition, see Phillips, A. S., *loc. cit*

VI.—REFERENCES.

[1] HILGARD, E. R. (1949). Human motives and the concept of self. *Am. Psychol.*, 4, pp. 374-384.

[2] JERSILD, A. T. (1952). *In Search of Self* New York . Bureau of Publications, Teachers' College, Columbia Univ.

[3] KERR, M (1952). *Personality and Confict in Jamaica.* Liverpool University Press.

[4] LECKY, P. (1945). *Self Consistency A Theory of Personality.* New York : Island Press.

[5] STAINES, J W. (1958). The self picture as a factor in the classroom. *Brit. J. Educ. Psychol.*, 28, pp 97-111.

(*Manuscript received 12th November*, 1962.)

VALUE OF THE ORAL RESPONSE IN BEGINNING READING: AN EXPERIMENTAL STUDY USING PROGRAMMED INSTRUCTION*

By JOHN D. McNEIL AND EVAN R. KEISLAR
(University of California, Los Angeles.)

SUMMARY. The question of the effect of vocalization in reading has not received a clear answer. In the present study, the procedures of programed instruction were used in testing the hypothesis that the saying of words and sentences out loud by beginners in reading will result in greater ability to recognize and understand written words and sentences. An auto-instructional program which taught forty words was administered to 182 non-readers in the kindergarten during a seventeen-day period. These children had been paired and assigned to oral and non-oral groups where they received identical programed instruction with the exception that those in the oral group were directed to 'say the words' instead of to ' look at the words.' An examination which called for silent reading, not oral reading, was given at the end of the last lesson The results supported the hypothesis that the oral response is valuable in learning to read.

I.—INTRODUCTION.

How important is it for beginners to say words out loud in learning to recognize and understand written words and sentences? In order to collect evidence regarding this perennial and fundamental question, procedures of programed instruction were used since they offer a high degree of reproducibility and control.

The psychological and educational literature has not given consistent answers to the value of oral responding in learning to read. Recognition and recall of objects have been found to be greater if the names of the objects are vocalized by the learner [2, 3, 6]. Jeffrey reported research which suggested that self-stimulation, the responding to one's own responses, helps in the overcoming of difficulty in learning to apply distinctive names.[5] One's own saying of a word might serve as a stimulus as well as a response. It is a stimulus in that the individual can hear and feel himself talk. " The feel of saying the word is part and parcel of the same pattern of speech stimulation."[4]

On the other hand, the observation that many poor readers verbalize has led some to the opinion that an oral response in learning to read is detrimental to reading achievement. Acting on this assumption, educators like McDade[7] and Buswell[1] sought favorable results from a method of beginning reading which required no oral responding and in which children were told not even to say the words to themselves. Learners were only to match words with objects, pictures, and actions.

The specific hypothesis of the present study was that kindergarten children taught by programed materials which required the saying of words and sentences out loud would achieve greater skills in reading than children taught by the same programed materials with no instruction to speak out loud.

---

* This research project was supported by a grant from the United States Office of Education, Department of Health, Education, and Welfare (Grant No. 1413)

II.—Design of the experiment.

The general plan of the study called for presenting an auto-instructional program, consisting of seventeen daily lessons, to a number of kindergarten children in small groups, half of whom were required to make oral responses and half of whom did not. At the end of the last lesson, the extent of learning was measured by a post-test. The presentation of the program to the small groups constituted replications of the experiment, providing evidence of reliability of the treatment and its effects.

Subjects.

The children used in the study had an average chronological age of 5 years and were the total kindergarten populations at three selected public elementary schools. One school represented an upper-middle class native-born white community; one school was located in a lower-middle class community, predominately of Japanese and Mexican backgrounds; and a third was located in a middle class community typical of those in most Southern California cities.

Before commencing the experiment, each child was given an intelligence test as well as a reading readiness test which measured his knowledge of letters, words to be taught in the program, and ability to recognize likeness in letter configurations. Children from the same classroom were paired on the basis of scores obtained from these tests and also upon their record of absenteeism. One member of each pair, randomly selected, was assigned to an oral group, the other member to a non-oral group. From these pairs thirteen oral and thirteen non-oral groups consisting of ten or less were constructed.

There were 231 children in the original sample. The subjects who were dropped from the study were: forty-three children absent more than three consecutive days, four children unable to speak English and two children identified by their teachers as hyperactive (emotionally disturbed). The remaining 182 children, therefore, made up the subjects for the study. These children had an average intelligence quotient of 107 with a range from 72–138 as measured by the Kuhlmann-Anderson Test of Intelligence.

III.—Method.

Programed Materials.

The instructional program consisted of seventeen daily lessons which took slightly over three weeks to administer. Each lesson was made up of approximately thirty-five questions or frames. Two sets of the programed materials, identical in all respects but one, were prepared. The non-oral set carried a taped commentary which directed the learner to " look at the word(s) "; the other set directed the learner to " say the word(s)." In order to assure that children in both groups were attending to the reading materials, all subjects were required to respond to problems throughout the program; for each problem, children were to select one of three choices offered by pressing an appropriate button. The oral group, therefore, made both oral and button-pressing responses; the non-oral group made only the button-pressing response.

Forty words were taught during the program both as independent words and as sentences. These words represented terms commonly found in the oral vocabulary of young children[8] and included both function words (i.e., and, with) and those with a direct referent (i.e., mother, girl). The printed words taught in this study were among those used in everyday speech by children. The study was not concerned with extending the child's vocabulary.

Examples of the 700 frames in the program are shown below :

	Purpose.	*Frame.*	*Commentary.*
(1)	To introduce a new word.	Children children woman dog	" Children. There are children in your room at school. The word at the top says children. (Say the word.) (Look at the word.) Which word below also says children ? "
(2)	To facilitate reading from left to right and to emphasize relations of a printed word to a pattern of language structure.	children run ...... fast girl boy	" (Say) (Look at) the words at the top. Which word below sounds right with the top words ? "
(3)	To maintain motivation as well as to test for comprehension.	the boy and girl can run picture picture picture A B C	" (Say) (Look at) the words at the top and find the picture that goes with the words."
(4)	To help the learner identify sentences which are appropriate answers to the questions he hears.	Billy Billy we go can eat runs fast	" What does Billy do when he is in a hurry ? "
(5)	To help learners rely on more than one cue in identifying words. (e.g., length instead of initial letter or configuration alone).	children cold chair	" Which word says children ? "
(6)	To give practice in recognizing and comprehending function words.	children ...... run we fast can	" Find the missing word."

The program aimed to develop four genaralizable skills. Beginning with the first lesson, efforts were made to teach children (1) to respond appropriately to new combinations of printed words, arriving at a complete thought or meaning, (2) to identify the particular word in a sentence which provides an answer to a posed question, (3) to respond to a variety of cues in discriminating among words, and (4) to read from left to right. To accomplish these four objectives, the following principles of program preparation were adopted :

(*a*) In order to get the learner to read from left to right, new words were introduced on the right of previously learned words (e.g., Frame (2) above).

(*b*) Pictures were rarely used in the program. The responses to printed words were prompted orally by the taped commentary. Occasionally, for motivational purposes, children would be asked to look at or say the words at the top and then find the picture from among three that " went with the word " (e.g., Frame (3)).

(*c*) The sequence of frames was built around reiteration of the words and the generalizable skills of word and sentence recognition. There was no narrative or thematic structure. Words and sentences were reviewed at unexpected times.

The overall difficulty level of the program (percentage of errors on first trial) was 33 per cent for the oral group and 38 per cent. for the non-oral group. Performance by individuals reflected a range in difficulty from 70 per cent. errors to 2 per cent. The average difficulty of the daily lessons ranged from 23 per cent. to 43 per cent. errors (oral group) and 23 per cent. to 51 per cent. (non-oral group).

The Apparatus.

The apparatus provided for instruction of ten or fewer children at a time. Each child sat in an individual laboratory-type cubicle equipped with headphones through which he heard the taped commentary. The cubicle also contained a microphone as well as response and confirmation panels. The response panel consisted of three buttons by which the child indicated his selection in answering the multiple-choice problems in reading. Confirmation of the correct answer was by lights ; a green light was shown when the learner made the correct response, and a red one when he was in error.

A single 6-ft. × 8-ft. screen at the front of the room was visible from each cubicle. On this screen, learners saw each daily sequence presented by filmstrip, while simultaneously hearing the accompanying taped commentary. The program advanced automatically at a paced schedule with 3-seconds allowed for responding to each frame. As each individual in the oral group made vocal responses into his microphone, he heard his own voice. Frequent monitoring of the individual oral responses through the microphone circuit gave evidence that subjects were responding orally as directed. When the non-oral group used the equipment, the microphones were deactivated. A record of each individual's multiple-choice performance for each frame was obtained by means of a punched paper tape.

The apparatus, therefore, combined advantages of both group teaching (i.e.. economy, control of word exposure time and pacing) and tutorial instruction (i.e , personal communication without peer group distraction and provision for individual response and confirmation).

The Testing Procedure.

The criterion test provided evidence that the learner had acquired knowledge of the words and reading skills previously mentioned. This post-test consisted of fifty-one multiple-choice items, each with three or four alternatives, similar in format to the frames in the program except that the items appeared in individual booklets, one item per page. Instructions for each item were given orally. No item in the post-test was identical to any frame in the program ; the context and the arrangement of alternatives had not been seen previously by the pupil. Twenty-four items required that the learner recognize a given word. The remaining items required the child to select the individual word or sentence which supplied the answer to an oral question, or an incomplete sentence, or which best described a picture.

Examples of the test items are as follows :

Objective.	*Instruction.*	*Test Item.*
He selects an individual word when it appears with those of (*a*) like and (*b*) unlike shape.	" Find the word that says dog."	(*a*) boy, dog, toy. (*b*) children, tree, dog.
He selects the word that will complete a sentence never seen before.	" Find the missing word."	a dog can play ...... a ball red, with, no, little

The test was administered in group situations, twelve children at a time, oral and non-oral together.

IV.—RESULTS AND CONCLUSIONS.

(1) *Overall Effect of Oral Response.*

The mean score for the oral group on the criterion test was 30·3 (SD=10·0) while the non-oral mean was 23·8 (SD=7·9). This superiority on the part of the oral responders was significant when based on the average difference between matched pairs (t=5·64, $p < \cdot 01$). Since the mean score for forty-seven uninstructed kindergarten children was only 12·7 (SD=3·3), both the oral and non-oral forms of the program appear to be effective in teaching.

The superiority of the oral response group is illustrated in Figure 1 where the results are presented for the three groups in terms of percentage of total possible scores. Since there was no correction for guessing, the dotted line indicates a chance score which would have been obtained by blind guessing. Using this line as a base, it may be noted that the effect of the oral response is considerable.

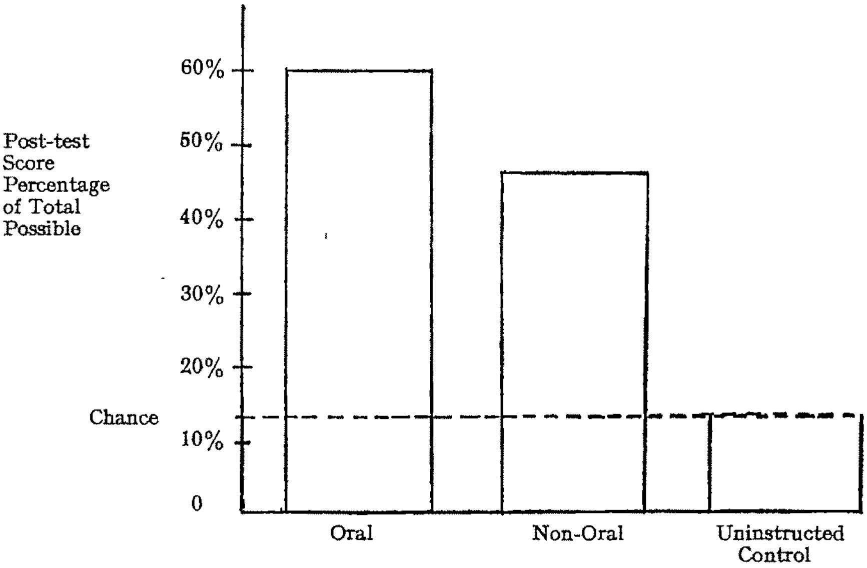


FIGURE 1

COMPARATIVE CONTRIBUTION OF ORAL AND NON-ORAL RESPONSE MODE IN READING ACHIEVEMENT.

Since the experiment was carried out with thirteen pairs of matched groups, we may also view the data non-parametrically. In each of the thirteen pairs of groups, the mean of the oral group was higher than the mean of the corresponding non-oral group. On the basis of a sign test, this difference is statistically significant ($p < \cdot 01$).

(2) *Value of Oral Responding at Different Levels of Ability.*

Although it was speculated that the oral response would be especially important for children of lower intelligence, the interaction effects were not large enough to be statistically significant. However, the direction of the results suggest the desirability of further inquiry regarding the special advantages of oral responding by those of lower intelligence. Table 1 below shows the mean scores by three levels of ability for both oral and non-oral responders.

TABLE 1

POST-TEST RESULTS FOR ORAL AND NON-ORAL SUBJECTS.

	Oral Written Post-test	Non-Oral Written Post-test
Low I.Q. (N=60)	M=25·2	M=17 0
Middle I.Q. (N=62)	M=31·8	M=24·1
High I.Q. (N=60)	M=33·7	M=30·3
Total N=182	M=30·3	M=23·8

(3) *Value of Oral Response in Motivating Children to Want to Read.*

In order to throw light upon the attitudinal effect of the program and the oral response, an interview was held with each child. Children from both oral and non-oral groups were individually asked whether they would prefer to continue with the program when given each of the following as an alternative: (1) listen to records; (2) listen to a story; (3) paint; or (4) play in the sandbox. Over 72 per cent. of the children in both groups preferred the reading activity to one or more of these other activities; 10 per cent. chose the reading activity in preference to all four of the other alternatives. Parenthetically, we can say that the introduction of a formal reading program was received with as much enthusiasm as currently popular activities in the kindergarten.

Although achievement in learning was not found to be associated with preference for continuing the reading activity, the data suggested that oral responding might have motivational properties for those of lower intellectual ability (I.Q. below 100). Preferential scores for oral and non-oral groups in middle and high I.Q. levels averaged 1·5, but preferential scores for the low I.Q. level were 1·6 and ·97 for the oral and non-oral groups, respectively.

(4) *Conclusion.*

The results show that oral responding by beginning readers facilitates the recognition and comprehension of printed words and sentences. Children who responded orally during the auto-instructional program were superior on a test which called for the application of a variety of reading skills in contexts different from those taught in the program itself. The oral response was of significant value even though the post-test called for silent reading, not oral responding, and none of the items of the post-test appeared in the training program.

As indicated earlier, one way to explain the mechanism involved in the greater achievement with overt vocalization is that the distinct stimulus of saying words aloud is required for mediation of the obtained effects.

The experiment gives no support for the belief that oral responding in learning to read has unfavorable effects. On the contrary, we must conclude that, although vocalization and poor reading on the part of the learner are often associated, the act of vocalization during instruction is an aid to the beginning reader in what for him is already a difficult reading situation. Without such vocalization, his performance would be worse.

V.—REFERENCES.

[1] BUSWELL, G. T. (1945). Non-Oral reading: a study of its use in the Chicago Public schools. Chicago, Illinois *Supplementary Educational Monograph, No.* 60

[2] CANTOR, G. N. (1955). Effects of three types of pre-training on discrimination learning in pre-school children. *J. Educ. Psychol.*, 49, 339-42

[3] DIETZE, D. (1955). The facilitating effect of words on discrimination and generalization. *J. Educ. Psychol.*, 50, 255-60.

[4] HORN, ERNEST, and CURTIS, JAMES F. (1949). Improvement of oral reading. *N.S.S.E. Forty-eighth Year Book, Part II*, p. 254.

[5] JEFFREY, WENDELL (1961). Perceptual learning in children. *Research Roundup on Children and Youth*, University of California Press.

[6] KURTZ, K. H., and HOVLAND, C. I. (1953). The effect of verbalization during observation of stimulus objects upon accuracy of recognition and recall. *J. Educ. Psychol.*, 45, 1957-64.

[7] McDADE, J. E. (1937). A hypothesis for non-oral reading: argument, experiment, and results *J. Educ. Res.*, 30, 489-503.

[8] STONE, C. R. (1956). Measuring difficulty of primary reading materials: a constructive criticism of Spache's measure. *Elem. Sch. J.*, LVII, 1, 36-41.

(*Manuscript received* 19*th November*, 1962)

THE INFLUENCE OF EARLY PUBERTY ON TEST PERFORMANCE AT AGE ELEVEN

BY

J. D. NISBET

(*Department of Education, University of Aberdeen*)

AND

R. ILLESLEY

(*Obstetric Medicine Research Unit, University of Aberdeen*)

SUMMARY. It has been suggested that the earlier onset of puberty among girls may now be affecting performance in tests at the age of 11 plus. This question was examined by the analysis of test scores at ages 7, 9, 11 and 13 from a year-group of girls numbering 1,385, for whom age at first menstruation and father's occupation were known The very early maturers proved to have a slightly higher average test score than the others, but they were superior at all ages and not only at 11. While the results at age 13 do not rule out the possibility of a later pubertal spurt, it can be concluded that early onset of puberty does not confer any advantage in selection tests at age 11.

I.—INTRODUCTION.

THE evidence for the secular trend towards early maturity has been summarized recently by Tanner (1961): for example, he reports that among girls in Norway, in 1840, the average age at menarche (or first menstruation) was 17½ years, whereas to-day it is about 13 years. The trend appears to have been maintained steadily over the whole of the last hundred years, and it is still continuing. (In earlier times, menarche seems to have occurred around the age of 14. Shakespeare's Juliet was 13: he devotes ten lines of Act I, Scene III, to emphasize that she is just " a fortnight and odd days " short of her fourteenth birthday.) The factors influencing the age of puberty remain uncertain; inadequate records and differences in mortality rates make research difficult.

This trend has aroused concern among those who are responsible for operating procedures of selection for secondary education. Girls reach puberty earlier than boys; and if there is a spurt in mental growth similar to the acceleration in physical growth at puberty, then as puberty occurs earlier, the spurt will be beginning to affect the performance of early maturing girls in tests at the age of 11 plus. Since the advantage of an early growth spurt is only temporary, this would reduce the validity of the selection procedure. Tanner (1962) has recently reported that early maturing girls are, indeed, scoring higher on average in selection tests than the late maturers. However, this finding does not allow for the possibility that the early maturers may have a consistent superiority at all ages over the late maturers.

Other indirect evidence would appear to be consistent with the hypothesis that scores at age 11 are beginning to be affected by the earlier onset of puberty, though here too other explanations of the findings are possible. The Scottish Mental Survey (1949), in which a test given to all the 11-year-old children in Scotland in 1932, was given again to the 11-year-old age-group in 1947, showed that the average score in 1947 was higher than in 1932. In 1932, the boys were slightly better on average than the girls. By 1947, the girls had not only caught up, but had passed the boys, and scored on average two points higher than the boys. Emmett (1950) in a similar comparison of age-groups in ten English areas over the period 1936-1947, also found that the average score of girls had increased relative to that of boys, though there was some indication that the girls' superiority had begun to fall away by the age of 13. Boyne and Clark (1959) found a similar increased superiority of girls over boys among Aberdeen school children over the period 1947-1957.

These findings would appear to be in accord with the hypothesis that a spurt in mental growth accompanying puberty is affecting performance at age 11. However, the data can be explained without assuming an adolescent spurt in mental growth. Differences between mean test scores of boys and girls may arise from a variety of causes. For example, girls are generally better than boys in verbal problems and boys are generally better in non-verbal and numerical problems. There is a tendency in many tests to use non-verbal material for the simpler items and verbal material for the more difficult items. Thus, if there is a slow general improvement in over-all performance (whether due to coaching or any other cause), this would also result in a trend towards girls' superiority as the whole age-group moved up into a type of item which tended to favour the girls.

For an adequate test of the hypothesis, a longitudinal study is necessary. An age-group must be tested at intervals from age 7 (or from any age at which it can be safely assumed that puberty has not yet affected test performance), up to at least age 13 (by which age it can be assumed that the influence of puberty, if it exists, will have begun to show an effect). Comparison of the average scores of the early maturers and the remainder provides a test of the hypothesis that there is a spurt in mental growth accompanying puberty, and will show whether earlier onset of puberty is affecting performance in tests at age 11.

II.—SUBJECTS, TESTS AND OTHER DATA.

The population studied was defined as those girls who were in the second year of secondary schools in Aberdeen—that is, around the age of 13 or over—and who were present on a day in December, 1960, on which one of our tests was given. For this year-group, the education authority records were searched to identify the earlier scores of these girls in the tests which had been given to the whole age-group at ages 7, 9 and 11, as part of the regular testing programme in the City of Aberdeen. On these earlier occasions, absentees were followed up to ensure that a test score was available for every pupil at each age; for lack of resources it was not practicable to follow up the absentees at age 13. Although all secondary schools, both local authority and independent schools, were included in the testing at age 13, earlier test scores were not available for most of the girls from two private schools attended by some 6 per cent. of the age-group of girls, as only a few of the girls at these two schools had taken the earlier tests. Migrants into the city also missed the earlier tests. In the end, complete test records from age 7 were traced for 85 per cent. of those tested at age 13. The sample is thus slightly biased towards the lower occupational groups.

The tests used were: at age 7, Moray House Picture Test; at age 9, Schonell's Essential Intelligence Test; at age 11, Moray House Verbal Reasoning Tests 57 and 58; and at age 13, Moray House Advanced Verbal Reasoning Test 10. Not all the sample sat the earlier tests at the same time: thus, at age 11, the large majority sat Moray House Tests 57 and 58, but a few sat the tests applied the previous year, Tests 55 and 56. This difference had to be ignored: age-allowances provide for differences in age at testing, and the Moray House Tests are intended to be, and generally are, comparable with one another. All the test scores were re-standardized on this sample at each age, to equate the means and standard deviations and to produce normal distributions in each test, thus eliminating differences in standardization between tests. Means of 100 and standard deviations of 15 were imposed in a conversion based on scores grouped in intervals of five points: interpolations within intervals resulted in means which worked out slightly below 100 and standard deviations below 15. The means and standard deviations of the original scores and of the scores after conversion are set out in Table 1.

TABLE 1

MEANS AND STANDARD DEVIATIONS BEFORE AND AFTER RE-STANDARDIZATION.

Test at Age—	Original Scores		Revised Scores	
	Mean	s.d.	Mean	s.d.
7	106·85	15·33	99·77	14·66
9	109·70	15·53	99·66	14·81
11 (1)	104·95	13·68	99·82	14·82
11 (2)	106·41	13·44	99·66	14·69
13	107·90	13·93	99·49	14·89

The next step was to find out age at menarche for each girl. The investigation was restricted to girls so that the age at first menstruation could be used as an indication of whether puberty came early or late. Since the sample was aged about 13 or over, menarche was a fairly recent event for many of them and it was, therefore, not necessary to rely on distant and possibly faulty memory. A considerable proportion had not yet reached menarche, but it was still possible to separate the early developers and the very early developers from the remainder. Existing medical records did not include age at menarche; but the School Health Service in the city undertook to interview individually all the girls in the sample to ask for this information. In three schools, the head preferred that the information should be got by writing to the mothers of the girls. This was done and practically all replied. Difficulties at this stage of the inquiry had been anticipated and particular attention was given to the method of approach. No one was scandalized, no protests were received, and information on age at menarche was obtained for 98·8 per cent. of those tested at age 13.

Information on social class was based on the answers to carefully designed questions on father's occupation. On this basis, pupils were graded into six groups on a modified version of the Registrar-General's Scale: I and II, professional and business; IIIa non-manual skilled workers; III skilled manual workers; IV semi-skilled manual workers; and V unskilled workers. When a pupil's description of the father's occupation was too vague to be classified, the pupil was asked, through the headmaster, for further information. The answer "Don't know (father's occupation)" was permitted, to allow for illegitimate children and other special cases. Only 2·3 per cent. of the total group could not be given a grading in one of the six categories and it was possible to establish from the detailed records of the Aberdeen Maternity Hospital that a considerable proportion of these were, in fact, illegitimate children.

III.—ANALYSIS OF RESULTS.

Complete test records—that is, scores on tests at ages 7, 9 and 13, and on two tests at age 11—were available for 1,402 girls. Information on menarche was lacking for seventeen of these, leaving a total of 1,385, from whom the following results are derived. The age range at the time of interview included six pupils who were under 13 years, and six over 14 years 6 months, but 91 per cent. of ages at interview were between 13 years 3 months and 14 years 2 months. The median age at menarche in the sample was 13 years 2 months. Menstruation began before the age of 11 in twenty-four cases, under 2 per cent.; before age 12 in 165, or 12 per cent.; and before age 13 in 40 per cent. of the group. Age at menarche was classified in three categories:

(i) The very early maturers, with menarche before the age of 12 years 3 months.

(ii) The early maturers, with menarche between the ages of 12 years 3 months and 12 years 11 months.

(iii) The remainder, among whom menarche occurred at or after the age of 13 years, or who, being already aged 13 years, had not yet reached menarche. (There were two girls not yet at menarche who were aged between 12 years 9 months and 12 years 11 months, and these were included in this third group).

This third group, therefore, included both the median group and the late maturers. It was not practicable to distinguish these. There was, in the sample, a group of 358 girls among whom menarche occurred after the age of 13 years 9 months. The mean scores of this group proved to be about five points of I.Q. below the average, but this was a spurious result arising from the nature of the sample. Where a sample is drawn from an educational year-group, there is a tendency for the brighter pupils to be younger than average and the duller pupils to be older. Younger pupils who were to prove late developers had not reached menarche at the time of interview. An attempt to establish a category of definite late developers from the available data would have missed these younger pupils and resulted in a group with a bias towards older age and lower ability. All pupils who were neither early nor very early maturers were, therefore, grouped together.

The same objection does not apply to the 'early' and 'very early maturers,' with whom menarche occurred before age 13 years or before 12 years 3 months, respectively, as these groups could be identified with certainty. It is thus possible to derive results which would test the hypothesis that the early maturers would show a spurt in mental growth around the onset of puberty, but not to obtain information on the effect with late maturers. It is, however, the early maturers who are concerned in the argument about the influence of puberty on performance in selection tests at the age of 11. The hypothesis under test is that, if there is a spurt in mental development with the onset of puberty which is beginning to affect performance in testing at age 11, this would appear as a sudden improvement of the very early maturers at the age of 11—an improvement, that is, relative to any possible superiority that they may previously have shown over the early maturers and the remainder at the age of 7 and 9 before any effect of puberty has begun to operate. By age 13, on this hypothesis, the lead of the very early maturers over the earlies would have been reduced, as the earlies, and perhaps some of the remainder, came to benefit likewise from this pubertal spurt.

The mean scores of the very early and early maturers and the remainder, which are set out in Table 2, do not show this pattern of results. The very early maturers are certainly superior at the age of 11, but this is not a sudden emergence of superiority with the onset of puberty. They are almost the same degree ahead at ages 7 and 9. In fact, the pattern of mean scores between ages 7 and 11 is, if anything, the reverse of what would be expected on the hypothesis of a pubertal spurt among very early maturers at age 11. Differences between the means of the menarche categories at any one age are generally significant (the ' very early ' at all ages over both ' early ' and ' remainder ' ; the ' early ' only at ages 7 and 13 over ' remainder ').

TABLE 2

MEAN TEST SCORES OF MENARCHE GROUPS.

Mean Test Score at Age—	Menarche		
	Very Early	Early	Remainder
7	103·5	100·7	98·3
9	102·8	99·4	98·8
11 (1)	102·9	99·9	98·9
11 (2)	102·6	99·8	98·7
13	103·3	100·2	98·0
N	246	300	839
Percentage entering Senior Secondary Schools	28	23	21

Beyond the age of 11, with the results at age 13, there is some suggestion that the early maturers are pushing ahead of the remainder. The differences between means at various ages for any one menarche category are not significant until we reach age 13 : at this age, the drop in mean score of the ' remainder ' category is significant. An alternative expression of this point is that the distance between the means of the ' very early ' and ' remainder ' categories increases significantly between 11 and 13. This may mark the beginning of a pubertal spurt among the very early maturers, though without a further testing at age 15 (now impracticable as many of the pupils have already left school) the data are inadequate to establish such a finding. However, it can be concluded from the present data that the earlier onset of puberty among girls is not seriously affecting test performance at age 11.

It should, perhaps, be emphasized that within each menarche group there is a wide range of test scores at each age, the very early, early and remainder categories all including pupils whose scores range from the highest to the lowest.

The relation of social class (assessed by father's occupation) to age at menarche and test score at various ages can be examined from the data. There is, of course, an association between social class and test score : the average of girls from Social Class I was between 114 and 117 on all the tests, and the average score of those from Social Class V was between 93 and 95. Again, there was a wide range of scores in each category. The superiority of the very early maturers cannot, however, be explained as merely the product of class differences. Even within any one social class, the very early maturers are consistently above the average for that class. The mean scores for each menarche category, separating the social class groups, are set out in full in Table 3.

TABLE 3

Mean Test Score and Menarche, By Social Class.

Social Class	Menarche Category	Mean Score at Age—					N
		7	9	11 (1)	11 (2)	13	
I/II	VE	115	116	114	115	116	(35)
	E	107	107	109	109	108	(47)
	R	108	107	108	107	108	(123)
IIIa	VE	107	106	105	105	106	(46)
	E	105	102	104	105	106	(37)
	R	97	101	103	103	101	(116)
III	VE	101	99	100	100	100	(111)
	E	101	99	99	99	99	(124)
	R	98	98	98	97	97	(372)
IV	VE	103	102	100	99	100	(22)
	E	94	96	96	96	97	(44)
	R	97	95	95	96	96	(98)
V	VE	97	97	99	98	98	(28)
	E	97	93	94	93	93	(43)
	R	93	94	92	92	92	(107)
Social class not stated : VE 4 ; E 5 ; R 23							(32)
					Total		1385

Sub-division by social class reduces the numbers in individual categories of the table, resulting in slightly larger chance fluctuations in the means at various ages. However, breaking down the data by social class produces no evidence to affect the conclusion derived from Table 2. Within each social class, the very early maturers generally have higher mean scores than the rest, but this appears at all ages and not merely with the onset of puberty.

Some association between social class and age at menarche might reasonably have been expected. When the investigation was being planned, the physiologically-minded among our advisers were inclined to predict that the early maturers would tend to come from the upper social groups, being taller, heavier and quicker to mature because of better nutrition. Some teachers, on the other hand, pointed out that girls from the poorer urban areas give the impression of earlier maturity in their appearance, dress, behaviour and interest in boys, and, in fact, they tend to marry earlier. Table 4, which presents the numbers and proportion in each menarche category for each social class, does not support either of these views : there is no marked association between age at menarche and social class in this sample. The proportion of very early and early maturers is nearly the same among daughters of professional and business men as among the daughters of unskilled labourers.

The data for Table 4 were drawn from all those for whom information on both social class and age at menarche was available, that is, 1,385 girls on whom the previous tables are based, less thirty-two for whom a social class grading was not available, plus 211 girls for whom test records were incomplete (mostly from private schools and social classes I and II). A parallel analysis of the 1,385 cases gives similar results.

TABLE 4

Distribution of Menarche by Social Class.

Menarche Category	Social Class					Totals
	I/II	IIIa	III	IV	V	
Very early	50 (16)	49 (21)	119 (18)	24 (13)	31 (16)	273 (17)
Early	67 (22)	44 (19)	131 (20)	48 (26)	46 (24)	336 (21)
Remainder..	189 (62)	135 (59)	404 (62)	110 (60)	117 (60)	955 (61)
Totals......	306 (100)	228 (100)	654 (100)	182 (100)	194 (100)	1564 (100)

(Percentages in brackets.)

Chi-square on this table is 8·6, with 8 degrees of freedom, and is not significant (P about 0·4).

The inter-correlations between the test scores at the various ages were as follows:

Between score at age 7 and score at age 9: +0·69
age 11 (1): +0·72
age 11 (2): +0·75
age 13: +0·74
Between score at age 9 and score at age 11 (1): +0·82
age 11 (2): +0·82
age 13: +0·82
Between score at age 11 (1) and score at age 11 (2): +0·93
age 13: +0·90
Between score at age 11 (2) and score at age 13: +0·91

IV.—Conclusions.

The results from this analysis of the test scores of 1,385 girls at ages 7, 9, 11, and 13, show that the anxiety that earlier onset of puberty is affecting performance in selection tests at age 11, is unfounded. On other aspects of the effect of puberty on mental development, the data do not permit any definite conclusions to be reached. The results at age 13 can be interpreted as indicating the beginning of a possible later spurt. It is also possible (some may say probable) that the onset of puberty affects aspects of mental development which are not measured by the tests used. Moreover, the investigation was concerned principally with a comparison of the early and very early maturers with the remainder, and the data were not appropriate for a study of the late maturers. Whatever the causes of the earlier onset of puberty, the evidence so far does not show any relationship between early onset of puberty and social class. Other aspects of this investigation, including family size, migration, type of secondary school attended, area of residence and birth weight, will be discussed in a subsequent paper.

ACKNOWLEDGEMENTS.—This study was begun in collaboration with David Hewitt, of the Department of Social Medicine, Oxford University, who was prevented by absence abroad from participating in the analysis of results. It will be obvious that the investigation involved the collaboration of a large team of workers, who assisted in the planning and organisation, in the administering and marking of tests, in the collection of data, in maintaining the records from which data were drawn, in searching these records, in coding and sorting the data of the investigation proper, and in providing information and assistance generally. Acknowledgement is due to the Health Visitors in Aberdeen under Dr. I. A. C. MacQueen, Medical Officer of Health; Dr. D. Younie, of the Aberdeen School Health Service; the headmasters, head-mistresses and teachers in Aberdeen secondary schools; Mr. W. Boyne of the Rowett Research Institute, Aberdeenshire; Mr. J. R. Clark, Director of Education, Aberdeen, and Mr. Johnson and Mr. Argo on his staff; Miss Ann Lamont and other Ed.B. students in the Department of Education, University of Aberdeen; and the clerical staff of the Obstetric Medicine Research Unit in the University of Aberdeen.

V.—REFERENCES.

BOYNE, A. W., and CLARK, J. R. (1959). Secular change in the intelligence of 11-year-old Aberdeen school children *Human Biology*, 31, 325-33.

EMMETT, W G. (1950) The trend of intelligence in certain districts of England. *Population Studies*, 3, 324-37.

TANNER, J. M. (1961). *Education and Physical Growth*. London : University of London Press. Chapter 7

TANNER, J. M. (1962). As reported in *Times Ed. Supp.*, Feb. 23, p. 355.

The Trend of Scottish Intelligence (1949) Publications of the Scottish Council for Research in Education, 30. London : Univ. of London Press.

(*Manuscript received 20th November*, 1962.)

THE MEASUREMENT OF LEARNING ABILITY

By G W. S. MACKAY AND P. E. VERNON
(University of London, Institute of Education)

SUMMARY.—Intelligence tests are often assumed to predict a child's ability to learn, although they are seldom based on situations that directly involve any learning. A battery of nine group tests was devised for junior school children, each based on the learning of an unfamiliar task, and yielding six scores for successive stages of practice. These were applied to 8-9 and 10-11 year pupils, along with a number of more conventional tests, in order to find whether they added appreciably to the predictions of achievement derived from the previous year's work. As a measure of learning, residual gain was used, i.e., final scores holding initial scores constant.

These residual gain measures were found to be largely specific in the younger group, though somewhat more consistent and more predictive in the older group However, the tests differed considerably in their validities, and further work might produce more useful ones. There are considerable difficulties in finding tasks relevant to school progress which will also cover a wide range of ability. A stronger case might be made for using final scores, after a period of instruction and practice on new tests, as predictors.

I.—INTRODUCTION.

THE main object of giving intelligence tests to children in school, or at an educational clinic, is to assess their capacity to make scholastic progress. This implies that the tests predict ability to learn ; and, indeed, several writers have defined intelligence as learning ability (Buckingham 1921, Dearborn 1921). If, however, we regard learning as changes in performance following instruction and practice, then, with Woodrow (1946), we "cannot simply state that improvement with practice is identical with some hypothesized intelligence." Much the same point is made by Simrall (1947), who adds that there is little or no evidence that gains due to practice on any one test are associated with gains on other tests ; i.e., there is no such thing as generalized learning ability. Intelligence is normally measured by tests which are given once only, and which provide no opportunity for practice or the acquisition of new mental skills. Hence, they bear very little relation to the amount of progress children are likely to make in their school work over a given period, say a school year. And yet they still correlate highly with future overall scholastic achievement. J. E. Anderson (1940) has pointed out the main explanation of this apparent paradox, namely that mental development is always cumulative. A child starts the year with a fund of information and mental skills, X, which are very similar to those involved in doing an intelligence test, I ; during the year he acquires further skills, Y. Now the correlation of I with final achievement, X+Y, is still quite large because r_{IX} is large, although the correlation r_{IY} may be very small or even zero.

Surely, it should be possible, then, to devise tests which actually involve learning in a controlled situation, and which might add to the predictions of later performance already covered by intelligence tests. Such tests should be of considerable value to the psychologist and/or the teacher. Furthermore, they should provide even better opportunities than current intelligence tests for observing and assessing children's methods of tackling difficulties, their attention, perseverance, etc.—qualities which are highly relevant to future success or failure.

II.—PREVIOUS INVESTIGATIONS.

A considerable amount of work on the factors involved in various types of learning has been carried out, though little of it has much bearing on scholastic achievement. Anastasi (1930, 1932), Thurstone (1938) and others, have demonstrated the existence of rote memory factors in the learning of paired associates, words, digits, etc.; for the most part, these seem to be fairly specific to the type of material and the conditions of presentation and recall. There is no evidence that they correlate with school learning, except in so far as they involve *g* or verbal ability. Vernon (1961) claims, in his review of the field, that logical memory or the learning of meaningful materials depends wholly on these familiar factors, i.e., that no special faculty of meaningful learning can be distinguished.

Another approach is represented by the investigations of Melton (1947) and Fleishman and Hempel (1954, 1955) into psychomotor learning, where measures of proficiency at different stages of practice were factorized, and changes in factor patterns were obtained as the learners became more proficient. In the early stages of practice at complex co-ordination or discrimination reaction tasks, there were considerable correlations with spatial, mechanical and other cognitive factors, suggesting that learners are tackling the task by means of their already established abilities. But at later stages there is less dependence on these familiar factors, and other motor factors become more prominent, or a factor specific to the particular task emerges. This explains why it is so difficult to predict ultimate motor skills either by paper-and-pencil aptitude tests or from success during the early stages of learning. A similar investigation of learning a short musical theme was published by Burroughs and Morris (1962).

Allison (1954, 1956) and Mollenkopf (1957), as part of a programme for selecting U.S. Navy personnel, designed performance tests involving the learning of mechanical tasks, in which subjects were instructed in assembling various models, and tested after each period of instruction. They found that these learning scores predicted achievement at the end of mechanical training at least as well as the Navy's Basic Battery of aptitude tests, and that they were capable of adding appreciably to the predictive efficiency of this Battery. Although no similar investigation has been carried out with scholastic materials, these results seem to offer some promise that observation and measurement of a pupil's progress in a learning situation might likewise improve on the prediction of his future achievement. G. R. Ortar (1960) was faced with the educational selection and guidance of Israeli children from very varied cultural and linguistic backgrounds. Finding that scores on standard tests were seriously affected by such differences, she introduced a period of coaching and, on retesting, obtained much improved validities in predicting school achievement. She did not, however, make use of measures of the gain brought about by coaching.

The most quoted studies of actual learning gains are those of Woodrow (1938, 1939), who practised two groups of subjects repeatedly on a number of simple tasks—cancellation, arithmetic, anagrams, length estimation, etc. Here also the factor content of the initial and final scores differed considerably, and there was little evidence of any general factor among the gain scores, nor of overlapping between the amount of improvement and intelligence. Nevertheless, gain scores on different types of task, e.g., numerical, perceptual, did show some positive correlation with *N* and *P* factors. One reason for Woodrow's negative results may have been that the tasks employed were of a rather simple kind, and Tilton (1953) claims considerably greater consistency in an investigation of school subject learning. Here, standardized achievement tests in eight subjects were given to fourth grade pupils, and repeated twenty months later. There were marked positive correlations, averaging 0·30 (or 0·51 when corrected for attenuation) between the gain scores, indicating a prominent general factor in school progress. Gains also correlated positively with initial achievement; i.e., those who were more able in any subject tended to progress better in it than those who were initially weak.

More recently, two large-scale investigations have been reported from the Educational Testing Service, Princeton, N.J.: Stake (1958) applied a dozen learning tasks, each involving several trials, to 240 seventh grade pupils. Both verbal and non-verbal, rote and meaningful, tasks were used, and an extensive battery of reference tests was added. Each learning task was scored for total errors or overall success, for slope of the learning curves, and for regularity of progress. The last set of measures yielded mostly negligible correlations, but the success measures correlated positively with intelligence and with scholastic achievement. On factorizing his complete battery, Stake concluded that there is no general learning ability, but a number of factors for different kinds of tests; also that there is little relation between slope or rate of learning and familiar psychological factors.

Allison's (1960) study was similar, but it included a greater variety of learning tasks (psychomotor, mechanical, rote learning and concept formation), which were given to 315 adult recruits. These were scored for average rate of progress, and for faster or slower progress in the second than in the first half of trials. The main factors among these measures included: (1) Conceptual learning; (2) Mechanical and motor learning; and (3) Rote learning. On comparing these with the factors established in a battery of thirty-nine reference tests, he observed, unlike Stake, that No. 1 was related to general verbal intelligence and number ability, No. 2 to a spatial-mechanical factor, and No. 3 to a rote memory factor.

It would seem then that learning progress is certainly not identical with intelligence, though there may be some positive overlapping when complex learning tasks are studied. There is little evidence, also, of a general learning ability, common to all types of task, though this too may be affected by the level of complexity of the tasks. Rather, learning measures tend to group under a number of rather ill-defined factors, according to the type of material and the conditions of learning and recall. Some of these group factors do seem to relate to ability factors, for example verbal learning correlating more highly with verbal ability, non-verbal learning with spatial ability, etc.

III.—THE PRESENT EXPERIMENT.

Our own small-scale experiment was conducted against this background with two groups of approximately thirty-six primary school children, aged 8–9 years and 10–11 years, being the complete unstreamed second and fourth years of a school in a good middle-class area. A study with these numbers can be regarded only as a first pilot exploration; but it may suggest ways of tackling a larger sample with more refined techniques if worthwhile results are indicated.

The aim was to devise a series of group learning tests which would approximate more closely to the normal school learning situation than those employed by Stake. Group tests were preferred, not only because the tests would inevitably be too lengthy for individual application, but also because children might well learn differently when individually instructed from the way they learn in class. Each test was completed in a school period; i.e., long-term learning and retention were not involved. All tests were given during January, 1961, by the first author who, as head master, had often conducted classes with these children. In addition, a series of conventional ' reference ' tests of verbal, numerical and spatial abilities and scholastic attainments was given, partly with the aim of exploring the factorial content of the new learning tests, partly in order to find whether the learning tests added significantly to conventional tests in predicting achievement. The final criterion of achievement consisted of a combination of the class teacher's estimates of progress made during the remainder of the school year, and standardized English and arithmetic tests.

An inherent difficulty in devising learning tests is that the more closely they parallel school learning, the more they are likely to be affected by previous experience and achievement. Another problem is to ensure that they will cover a wide enough range of ability to be useful throughout the junior school. Ideally, a dull 8 or even 7-year-old should be able to make a start with them before he receives practice ; while at the same time a bright 11-year-old should still be improving at the end of practice. A description of the tests will show how these difficulties were tackled. Details may be found in Mackay (1962).

(1) *Language Learning Test.* This test attempted to teach the children a new language that they could translate from or into. Esperanto was used because of its inherent logical structure, and because it seemed unlikely that any children would be familiar with it. The test was a printed one and was tackled without any time limit, although it became clear that after about 40 minutes, the children had reached their limit. The initial instructions were not standardised ; they were designed to reduce anxiety by giving the impression that this was a normal lesson, not a formal test. But once the nature of the task had been pointed out, the children carried on by themselves from printed instructions.

The first exercise showed a few easy English words with their Esperanto equivalents. Those carrying the definite article were prefixed with La, those with the indefinite article merely had the equivalent without La. This exercise expected the children to grasp this rule, and they had to use it in translating the given words. The rule was then carried forward to the next piece of information, and so on until sentences could be translated using the rules governing the pronouns, verb tenses, plurals, agreement of adjectives with nouns, etc. Finally, a complete piece of translation could be, and was, attempted by the children. In order to give an initial, final, and total score, the first third, last third, and total number of answers were marked separately.

(2) *Number Substitution Learning.* This test was similar in structure to the Language Learning. A simple form of algebra unfamiliar to the children was used. Letters were given equivalent numbers, and after some straightforward interpolation, additional rules were supplied involving addition, subtraction, multiplication and division, which, although familiar, to the children, did depend upon previous experience. It seemed impossible to deal with numbers in any complex way without assuming a knowledge of these processes. The rules, however, had to be applied in a way that was not familiar, and it was the interpretation and progressive learning of the rules that was primarily involved.

(3) *Verbal Instructions Learning.* This test was given six trials, as was every subsequent test. The children were provided with six printed sheets, each sheet containing eight rows, each row consisting of two numbers and two letters. The tester read the instructions for all eight rows : e.g., put a ring round number seven and cross out letter A. After hearing the complete list, the children answered as many items as they could on their first paper. The full instructions were read again, and a second attempt was made with a new sheet. This was done six times.

(4) *Verbal Information Learning.* The above test involved rather meaningless material. This test presented a kind of learning more similar to what children meet at school. Twenty-five pieces of information were read out; these were simple, but unlikely to be known by any of them: e.g., Bach died in 1750; the first five letters on an English typewriter are QWERT. After a first reading of the complete list, the children answered twenty-five questions based directly on the information. Six trials were given, and the order of asking the questions (*not* of reading the statements) was altered each time.

(5) *Word Memory.* A list of sixteen three or four-letter words was read out to the children, and they were given two minutes to write down as many as they could remember. (In practice they were allowed as much time as they needed). This was repeated six times.

(6) *Perceptual Detail Learning.* The above tests covered the kind of material, the children learn in the verbal and number aspects of their school work. This test and the two following attempted to see how they learn using visual cues. The children were briefly shown a large card with twelve circles representing faces Each face was incomplete in some detail of eyebrows, eyes, nose, and mouth. No other features were shown, and those that were shown used only dots or straight lines; no drawing skill was involved. Having looked at the twelve faces for eight seconds, they then completed a sheet containing twelve blank circles. This test was repeated six times.

(7) *Geometrical Form Learning.* This test presented a large, rather complex, geometrical figure, built up from squares, circles, triangles, diagonals, etc., twenty distinct points in all. After looking at the figure for eight seconds, the children attempted to draw this figure. Again, six trials were given.

(8) *Symbol Maze Learning.* Some kinaesthetic learning was added to the learning of visual patterns. Each child had a printed maze, and the class was shown eighteen symbols that could *not* be crossed on the maze. After seeing the forbidden symbols for eight seconds, the maze was attempted. Six trials were involved.

(9) *Group Concept Formation Test.* Although only one test of this type was included in the battery, we feel that there is much to be learned from a battery of tests exclusively of the concept-formation type, some being individual. Bruner, Piaget, Annett (1959) and others have indicated the importance of this aspect of learning, and further studies are clearly needed. This test involved the linking of nonsense syllables with pictures. The first example showed a horse—TIX, a dog—TIX, a man—FON, and a woman—FON. The children were then shown (*a*) a baby and (*b*) a cat, and asked to write the appropriate syllables after each. There were fourteen trials in all, some being repeats, with the complexity of the concepts increasing. Concepts in the middle range included things that fly, things that make a sound, and people in uniform; and at the end of the scale—things nocturnal, physical action, and modes of transport. In the older group at least it was clear that children had benefited from experience with the early, easy, concepts when tackling the later, difficult ones.

Reference Tests' The remaining tests in the battery were as follows:

(10) A forty-item, multiple-choice, vocabulary test.
(11) A reading comprehension test, similar in form to the Watts-Vernon test.
(12) Schonell's Essential Mechanical Arithmetic.
(13) Schonell's Essential Problem Arithmetic.
(14) A twenty-item number series test.

(15) Digit memory, fourteen items ranging from three to nine digits; this was given orally and responses were written.
(16) A spatial test involving the copying of printed figures onto a dotted space.
(17) Goodenough Draw-a-Man test.
(18) Bender-Gestalt, given as a group test with each figure shown on a large card.
(19) Overall school attainment had already been examined in 1960, and similar tests were given in the Summer of 1961 to provide a criterion of later achievement.

IV.—SCORING.

Mollenkopf and Allison seem to have used the total scores on all trials as their measure of learning. While these obviously reflect improvement, they depend very largely also on initial ability. Woodrow, on the other hand, took final-minus-initial scores, that is the absolute gain. The disadvantage here is that progress on most tasks is unlikely to be linear; thus, there can be no guarantee that the units are equal at different parts of the scale. For example, it might be easier for those who start with a score of 10 to finish with a score of 30 than for those who start at 40 to finish with a score of 50. Stake plotted individual learning curves for each task, but eventually employed three measures for each curve, namely total success, slope or curvature (i.e., a measure of gain), and goodness of fit (i.e., regularity of progress). Since the latter measures yielded few significant correlations with any other tests, we decided to study total score, initial score in the first two trials and final score in the last two trials; (the first and last would have yielded bigger differences, but would also be likely to be poorer in reliability). However, the amount of progress was determined by correlational analysis, as follows, not by subtracting initial from final scores.

Manning and DuBois (1958) argue that the most appropriate measure of learning is "that portion of the measure of final status which is not correlated with initial status." Thus, if we denote initial (or early) performance as A, final performance as Z, and compare these with some criterion, C, or with any other test, then the correlation of effective gain with C will be:*

$$r_{(Z-A)} = r_{ZC} - r_{AC}r_{AZ}$$

Manning and DuBois do not indicate how one should correlate gain scores on one test with those on another. But we can extend the above formula to Tests 1 and 2 as follows:

$$r_{(Z_1-A_1)(Z_2-A_2)} = r_{Z_1Z_2} + r_{A_1A_2}r_{A_1Z_1}r_{A_2Z_2} - r_{A_1Z_2}r_{A_1Z_1} - r_{A_2Z_1}r_{A_2Z_2}$$

Each of the nine learning tests yielded a total, an initial and final score in each of the two groups. With nine reference tests, 1960 attainment, 1961 estimates and 1961 attainment, this means thirty-nine variables in each group. The 1,482 product-moment correlations were kindly calculated by the London University Computer Unit,† and the gain score r's were derived from these by hand, according to the above formulae. The correlations of each measure with the two 1961 (criterion) variables were generally very similar; hence they were combined by Correlation of Sums formula. The resulting validities are shown in Table 1.

---

* Presumably this expression should be corrected, as in the ordinary formula for partial correlation, by being divided by $\surd(1-r^2_{AC})\,(1-r^2_{AZ})$. However, it seemed better to omit this step, since it resulted in boosting many of the Gain score correlations (because r_{AZ} was usually very large). In other words, we followed the same procedure as in calculating the residuals in a centroid factor analysis.

† These may be borrowed on request to the authors.

TABLE 1

CORRELATIONS OF LEARNING TEST GAIN AND TOTAL SCORES AND REFERENCE TESTS WITH ATTAINMENT CRITERION.

Learning Tests	Gain Scores		Total Scores		Reference Tests	8–9	10–11
	8–9	10–11	8–9	10–11			
Language Learning	·163	·201	·810	·598	Vocabulary	·567	·476
Number Substitut'n	·011	·547	·655	·787	Reading Compr. ..	·595	·666
Verbal Instructions	·256	·206	·563	·189	Mechanical Arith.	·777	·812
Verbal Information	·199	·203	·762	·688	Problem Arith. .	·692	·657
Word Memory	·112	·276	·770	·296	Number Series ..	·639	·783
Perceptual Detail..	·306	– 042	·517	·198	Digit Memory ..	·428	·258
Geometrical Form	·134	·361	·317	·102	Spatial Test	·599	·393
Symbol Maze	·150	·332	·465	·367	Goodenough Draw.	·382	–·117
Concept Formation	·065	·251	·535	·574	Bender-Gestalt ..	·311	·359
					1960 Attainment.	·870	·830

V.—RESULTS.

It may be seen that the validities of the Total scores are similar to those of the Reference tests, though rather variable. None is as predictive as all-round attainment at the beginning of the year. By contrast, the Gain score validities are low; yet all but one are positive. Note that there is no consistency here between the two age groups, suggesting that the tests with near-zero validities are those which were too difficult for one group to register much progress (e.g., Number Substitution for 8–9 years), or too easy for the other (Perceptual Detail for the 10–11 years). In any case the Standard Errors of correlations based on thirty-six cases are so large that inferences as to the worth of particular tests would be unsafe. For the same reason, neither multiple regression analysis nor factor analysis seemed likely to help in elucidating the results. Instead, we condensed the correlation matrices by Correlation of Sums formula to yield the correlations among groups of similar tests. These are shown in Table 2 (8–9 year above, and 10–11 year below, the diagonal).

TABLE 2

INTER-CORRELATIONS AMONG COMBINED GROUPS OF VARIABLES.

	No. of tests	Verb. Gain	Non-Verb. Gain	Verb. Total	Non-Verb. Total	Verb. Abil.	No. Abil.	Spat. Abil.	1960 Att.	Criterion
Verbal Gain	5		·131	·388	·264	·108	·265	·109	·289	·318
Non-Verbal Gain ..	4	·491		·219	·617	·092	·395	·260	·369	·272
Verbal Total	5	·595	·477		·590	·562	·725	·367	·796	·893
Non-V. Total	4	·411	·670	·616		·353	·599	·453	·630	·662
Verbal Ability . .	2	·371	·393	539	·402		·417	·501	·670	·625
Number Ability .	4	·622	·569	·789	·682	·608		·452	·752	·795
Spatial Ability	3	·208	·450	·219	·493	·222	·396		526	·526
1960 Attain.	1	·442	·419	778	·501	·700	·840	·315		·870
1961 Criterion	2	·512	·399	·702	·486	·621	·783	·285	·830	

All the gain score correlations with other variables are noticeably lower in the 8–9 year than the 10–11 year group, indicating poorer reliability and greater specificity among younger children. Even the combined Verbal Gain measures (Nos. 1-5) correlate only 0·131 with the combined Non-verbal measures (Nos. 6–9), as against 0·491 in the older group. In contrast the two sets of Total scores show very similar relations to other variables in the two groups (except to the non-verbal reference tests; the Goodenough, Bender and Space test had very little in common among 10–11 year olds). Thus, overall performance at the new tests is quite predictive in the second year of the junior school, but children of this age do not show as consistent an ability to make steady progress at unfamiliar tasks.

Secondly, it may be noted that Gain scores tend to correlate more highly with Totals on the same tests than with any other variables. Children with good ability in any task do tend to progress more rapidly at this or similar tasks.* The same point is brought out by the *mean* correlations of the nine Gain scores with:

	8–9	10–11
Other Gain scores	·051	·141
Own Total scores	·375	·472
Other Totals	·113	·174
Reference tests and other variables	·111	·215

It would follow that rate of progress on any one task will not tell the teacher or psychologist much about performance on any other task, particularly with younger children. At the same time, Woodrow, Stake and other writers are not justified in regarding learning ability as almost wholly specific. For example, the mean inter-correlation of 0·141 among 10–11 year olds implies that our battery of Gain scores would correlate 0·60 with another similar battery; and doubtless better tests could be devised. Moreover, a certain amount of patterning among the Gain scores may be observed in Table 2. In both groups the Non-Verbal Gains correlate more highly than do the Verbal Gains with the spatial-pictorial reference tests. However, the Totals for the learning tests show this patterning more consistently.

Prediction of Attainment. The matrices were now further condensed by combining all Gain scores, all Totals, and all Reference tests (but omitting the Perceptual Detail Learning and Goodenough in the 10–11 group). Multiple correlations with the 1961 criterion were calculated, and the beta weights were as follows:

	8–9	10–11
Gain scores combined	—·112	·093
Total scores combined	·546	—·003
Reference tests	·231	·198
1960 Attainment	·285	·613
Multiple R	·939	·845

---

* This might be expected from the derivation of the Gain scores, since the correlation of Gain with Total = $r_{ZT} - r_{AT}r_{AZ}$ and this expression can be zero or negative only if r_{AT} exceeds r_{ZT}. In the light of Fleishman's work, it seems unlikely that a person's initial performance could more closely resemble his overall success than does his final performance.

The results differ markedly for the two groups. In the 10–11's, the previous year's attainment is much the best predictor. The learning test Total adds nothing, but Reference tests cover some more ground, and Gain scores make a very small additional contibution. In the younger group the Gain scores add nothing to the other predictors ; indeed, their beta weight is negative. Previous year's attainment is somewhat less diagnostic than at 10–11, but both the Learning Test Totals and the Reference tests do contribute appreciably.

One other possibility is that the Final scores on the learning tests would not only be much more convenient than any form of Gain score but might also be the most predictive type of score. In so far as both the Total score and rate of progress must depend to some extent on the rather chancy initial reaction to an unfamiliar situation, it might be argued that the performance reached after a certain amount of practice at the new task would be a better indicator of scholastic capabilities. This is borne out among the 10–11 year pupils, the correlation of combined Final scores with the criterion being 0·705, which is higher than 0·661 for the combined Total scores. Note, however, from Table 2 that it is still below the validities of the combined Reference tests and of 1960 attainment, though it could probably be considerably improved if the more promising learning tests only were used. In the younger group the Final scores all give lower correlations than the Total scores with the criterion, though here too, some of the tests seem promising, e.g., Language Learning 0·797, Verbal Information 0·719.

VI.—CONCLUSIONS.

Our search for tests involving actual learning which might be more predictive of future scholastic achievement than 'cross-sectional' tests has proved somewhat disappointing, particularly with children in the second junior year, though more successful among older children. Apparently, such tests need to be graded carefully to the level of ability of the particular group of pupils, otherwise the amount of progress made in a limited time is highly unreliable and specific. However, it is not true, as some writers have maintained, that there is no consistency. When Residual Gain is measured, a considerable general learning factor appeared among the older children, and there was some indication of group factors for learning different types of materials—verbal, numerical, non-verbal, etc. Each Residual Gain score also correlated relatively highly with overall success at that particular task.

The predictive value of the various learning tests that were tried out varied widely, yielding the expectation that a battery of successful ones could be devised; though much more exploration with larger numbers of cases is needed. Finally, it would appear that rate of progress or gain scores possess no inherent advantage. It is the final scores achieved after instruction and practice at some unfamiliar task that yield the highest validities, and seem most likely to add to the predictive value of of conventional tests that are applied without such practice.

Useful qualitative observations can be made during the tests of pupils' methods of work, concentration, etc. However, the outlook for individual learning tests which could be applied in the educational clinic does not appear very promising, both because these would have to be very lengthy to allow adequate practice, and because a child's reaction to the learning situation in the clinic setting might differ considerably from his reaction in the classroom.

V.—REFERENCES.

ALLISON, R. B. (1954). *Learning Measures as Predictors of Success in Torpedoman's Mates School.* Princeton, N.J. · Educational Testing Service, ONR Project 151-113.

ALLISON, R. B. (1956). *Learning Measures as Predictors of Success in Pipefitter and Metalsmith Schools.* Princeton, N J. : Educational Testing Service, ONR Project 151-113.

ALLISON, R. B. (1960). *Learning Parameters and Human Abilities.* Princeton, N.J. : Educational Testing Service, Technical Report.

ANASTASI, A. (1930). A group factor in immediate memory. *Arch. Psychol.*, No. 120.

ANASTASI, A. (1932). Further studies on the memory factor. *Arch. Psychol.*, No. 142.

ANDERSON, J. E. (1940). The prediction of terminal intelligence from infant and pre-school tests. *Yrbk. Nat. Soc. Stud. Educ.*, 39 (I), 385-403.

ANNETT, M. (1959). The classification of instances of four common class concepts by children and adults. *Brit J. Educ. Psychol.*, 29, 223-236.

BUCKINGHAM, B. R. (1921). Intelligence and its measurement. *J. Educ. Psychol.*, 12, 271-275.

BURROUGHS, G. E. R., and MORRIS, J. N. (1962). Factors involved in learning a simple musical theme. *Brit. J. Educ. Psychol*, 32, 18-28.

DEARBORN, W. F. (1921). Intelligence and its measurement. *J. Educ. Psychol.*, 12, 210-212.

FLEISHMAN, E. A , and HEMPEL, W. E. (1954). Changes in factor structure of a complex psychomotor test as a function of practice. *Psychometrika*, 19, 239-252.

FLEISHMAN, E. A., and HEMPEL, W. E. (1955). The relation between abilities and improvement with practice in a visual discrimination reaction task. *J. Exper. Psychol.*, 49, 301-312.

MACKAY, G. W. S. (1962). *A Study of the Prediction of Learning Ability in the Primary School.* M A. Thesis, University of London Library.

MANNING, W. H., and DUBOIS, P. H. (1958). Gains in proficiency as a criterion in test validation. *J. Appl. Psychol.*, 42, 191-194.

MELTON, A. W. (1947) *Apparatus Tests.* Army Air Forces Aviat. Psychol. Prog. Res. Rep No 4. Washington, D.C. : U.S. Government Printing Office.

MOLLENKOPF, W. G. (1957). *The Gear Assembly Test Description and Preliminary Results.* Princeton, N.J. : Educational Testing Service. ONR Project 151-113.

ORTAR, G R. (1960) Improving test validity by coaching. *Educ. Res.*, 2, 137-142.

SIMRALL, D. (1947). Intelligence and the ability to learn. *J. Psychol.*, 23, 27-43.

THURSTONE, L. L. (1938). Primary mental abilities. *Psychometr. Monogr.*, No. 1.

VERNON, P. E. (1961). *The Structure of Human Abilities* (2nd edit.). London: Methuen.

WOODROW, H. (1938). The relation between abilities and improvement with practice. *J. Educ. Psychol.*, 29, 215-230.

WOODROW, H. (1939). Factors in improvement with practice. *J. Psychol.*, 7. 55-70.

(*Manuscript received 22nd November*, 1962.)

CRITICAL COMMENTS

TEACHING EDUCATIONAL PSYCHOLOGY IN TRAINING COLLEGES*

The Report of a Joint Working Party of the British Psychological Society and the Association of Teachers in Colleges and Department of Education set up in 1959 to study the teaching of psychology to training college students and to make recommendations.

I

By T. H. B. HOLLINS
(James Graham College, Leeds.)

The Committee of Psychologists of the B.P.S. and the A.T.C.D.E. has produced a valuable report. It consists mainly of an analysis of the results of a questionnaire, on the teaching of psychology, sent to colleges in January, 1960, and recommendations for the future. The tone is critical without being damning and the suggestions constructive ; they are not too idealistic to be practicable. The report is short and cogent ; it is remarkably free from the hazy, high-sounding generalizations which afflict so much writing in Education, particularly from committees.

The picture presented here of the teaching of educational psychology in the colleges, at the end of the two-year course, is an unfavourable one. First, the proportion of unqualified lecturers teaching psychology is too big—39 per cent. (I shall speak in the present tense, as it is doubtful if the position has changed markedly since 1960) : and the position is worse than this figure indicates because, of course, some colleges have a number of specialists, more than their ' share,' and too many colleges have none at all. In case it is argued that qualifications are unimportant, the report shows, from an analysis of examination papers in Education, set in 1959, of all the A.T.Os., that the teaching is unsatisfactory ; some of it is out-of-date, e g., questions on McDougall's list of instincts ; most of it does not take account of recent work, e.g., there were very few questions on Piaget's contributions ; and there were too many questions which could be answered merely by describing school experiences. In short, " there was little evidence . . . that students were expected to possess specific knowledge in educational psychology as they might be expected to possess it in geography or physics or English." In view of these conclusions, it is cold comfort to find that " the allocation of time to the teaching of educational psychology is reasonable " ; it is about 10–15 per cent. of the whole college course, and about 40 per cent. of the Education course.

The picture would be even blacker, I think, if the analysis had gone further. A closer examination of the qualifications of specialists in educational psychology would reveal that too many of them are inadequate. Too many diploma courses allow students to qualify by theses describing children's behaviour on a basis of inadequate theory and a lack of rigorous experimental design with appropriate statistics. Wiseman† pointed out some of these weaknesses in 1959 ; he analysed theses and papers in educational psychology and showed how few of them were concerned with the fundamental theory of learning and child development, and how strong was the bias towards descriptive studies of behaviour. Again, I am sure the committee would have been saddened by an analysis of the *answers* to the 1959 questions, instead of the *questions*. In the first place, in my experience, the average student, answering a paper in which he has to choose four questions out of ten, avoids any question in which he may be forced to produce precise information. Secondly, he tends to give the same answer to the four he does choose, e.g., on treating maladjustment—affection, security and play ; on learning—the essentials are affection, security and play , in fact, if the rubric were to forbid the mention of ' play ' few would be able to answer at all, at least, amongst students training for the primary school.

---

* Copies of the Report are available from the offices of the B.P.S. at Tavistock House South, Tavistock Square, W.C.1., or the A.T.C.D.E., 151, Gower Street, W.C.1., price 5s. (post free).

† WISEMAN, S. (1959). Trends in educational psychology. *Brit. J. Educ. Psychol.*, 29, 128-135.

Some excuse can be found for the shortcomings of the teaching if one considers such factors as the small amount of time available in the two-year course, the demands made on lecturers in most colleges, which prevented adequate study, the poor intellectual calibre of some of the students (and this report remarks on our tendency to spend a disproportionate amount of our effort on the near-failures), and the gaps and weaknesses in psychological knowledge. A worthwhile psychology text is difficult (particularly to women students, whose experience of mathematics has frightened them off any information presented in tables) and if students are getting nothing from their reading (or, worse still, learning off the text by rote) we are forced to increase the practical work (in itself, an essential approach to psychology and children) to the extent of leaving out most of the theory. Brearley,* in 1957, gives an outline of a practical course for ' primary ' students and she concludes, realistically, no doubt, " Comparatively few students are ready to understand fully the tenets of the major schools of psychology, but we have greatly failed if at least some of them are not ready to study them when they leave." Collier,† in 1960, working with men students, reports " A difficulty is getting students to read effectively . . . ; it is the *quality* of reading that troubles us " and finds that theory and practice only come together in ' Study-Practice ' ; he makes a good case for a variety of methods, mainly practical, in this article, and some of his ideas are incorporated in a valuable section on methods in this report. As for psychology itself, the report rightly warns us to make " a cautious use of psychological theory " ; but this is putting it mildly. We have some certainties in the field of abilities, but in the field of learning, and personality structure and development, which are central to the teaching situation, we are among the " Two-and-Seventy jarring Sects. " In addition, nearly all the terms we wish to use in educational psychology have built-in value judgments, which lower their scientific value ; the report regrets that so few questions are set on ' motivation,' but maybe the lecturers, after reading a few pages of Peters' *Concept of Motivation*, have shown a wise discretion. Of course, the study of psychology can sometimes be as good an education as the study of literature ; both will show the complexity of life and the unlikelihood of finding easy solutions to problems ; both might produce a desire to solve the problems, nevertheless , but few student-teachers are likely to reach this level of sophistication. Some lecturers, too, become tired of keeping up with what seem merely fashions in psychology.

However, signs of discouragement are, in fact, seldom observed. The comparative neglect of many areas of educational psychology in the training colleges, up to 1960, when the three-year course came in, was due, not to pessimism about the value of psychology, but to an optimistic concentration on an educational policy which took up such psychological contributions as reinforced the policy and neglected the rest. This policy was, and is, to make the schools child-centred. A training college educationist, especially in primary work, would see the task as persuading the 18-year old student to change from the concept of learning she would have gained in her recent education, i.e , close study of a subject, to a concept of individual children learning in a controlled but permissive environment where social and emotional learning are as important as intellectual.

In this atmosphere, gentle but insistent pressure is put on both student and child to adjust to the norm of the ' happy individual ' which seems to be the ideal of the clinical psychologist. McDougall's theories seem to give psychological respectability to this aim—all the instincts and emotions working together, in the end, under the happy jurisdiction of the master-sentiment Freud will not do ; sex and aggression seem always about to overcome reason and civilization ; or, if repression is complete, neurosis seems inevitable ; and there is a distasteful pessimism in his dictum ' The function of education is to inhibit, suppress and forbid.' I am laughing only at the excesses of the movement and deploring only what sometimes amounts to indoctrination, because, of course, one must have admiration for a policy which has helped to produce our fine infant schools and which, even if it has not yet caught up fully with Piaget, gives our students a practical understanding of such important psychological concepts as individual differences, reading-readiness and 'learning through activity and experience'; and which produces a flow of colourful and interesting books for teaching reading and number, with efficient methods for using them.

---

* Brearley, M. (1957). The students' understanding of children. *Educ. for Teach.*, No. 42, 13-18.

† Collier, K. G. (1960). The teaching of psychology in training colleges. *Brit. J. Educ. Psychol.*, 30, 103-108

What one wants, of course, is to retain the benefits of this crusading educational outlook and to add to it the benefits which come from the more dispassionate scientific outlook which takes a cool look at the evidence and believes in experiment. It is here that a deeper, more searching, study of psychology in the training colleges, led by a psychologist properly trained in experimental procedures, would be of value. Luckily, conditions for this to come about are more favourable than in 1960. The introduction of the three-year course will give students more time to mature and to gain a better understanding of psychology. Even more important, there is a determined effort, in the institutes and colleges, to ensure that members of staff are allowed more time for reading and research. The institutes have, for many years now, provided advanced teaching in psychology, but now many more lecturers have the time and opportunity to attend them. Then, the expansion of the colleges has meant that the large college is now the rule, rather than the exception, and this means that the appointment of a qualified psychologist in nearly every education department is a possibility. Further, as the report points out in the chapter on "The Standard of Students," the colleges "continue to attract an appreciable proportion of students who have high natural endowments" and we know that, owing to the inability of the universities to expand quickly enough, we are likely to get an increasing proportion of highly intelligent students, although the range of ability is likely to remain a wide one. It is against this background that the recommendations of the report appear particularly relevant and practicable. All of them are important, but the following seem to me the four which would contribute most to a rapid development of the sound teaching of psychology and to the growth of a scientific outlook amongst the better teachers:

(1) It is desirable that at least one member of the staff should have a degree in psychology or its equivalent in a Bachelor's, Master's or Ph.D. degree in Education, in which the main content of the course is in the field of educational psychology.

(2) . . . the teaching in educational psychology, being a co-operative task in most colleges, requires that members of staff concerned with the subject should work as a team and the fullest use should be made of the specialist knowledge of the lecturer best qualified in educational psychology.

(3) More use should be made of the facilities which exist for obtaining further qualifications through full-time and part-time study. (The facilities are summarized in Appendix 6.)

(4) We press strongly for the inclusion of more advanced optional courses in educational psychology for the abler students.

It is particularly desirable that advanced courses in psychology should be made available to the more gifted students. Although the chief reasons for this are mainly of the idealistic kind, already mentioned, it should not be overlooked that the training colleges are hoping to have their certificate recognized in some way by the universities and even to have their better students working on degree-standard courses in Education. This would link up with the growing demand to have Education, and particularly educational psychology, as a first-degree subject in universities.

Of course, no startling advances can be looked for immediately. The chief obstacles to the development of psychology as a science lie outside the ranks of psychologists and training-college lecturers. Some of these are—the indifference, or even contempt, of universities, public schools and grammar schools towards psychology, which means it will continue to be starved of its proper share of the top-level of ability, and of funds for research, for some years yet; the large classes in schools, which force the teacher to be unduly concerned with class management and keep him suspicious of a science which thrives on the variations amongst individuals; and the unenlightened teaching of mathematics and the physical sciences in schools, particularly to girls, which results in the inability of most students to study psychology beyond the anecdotal stage. Nevertheless, this report makes useful suggestions for another step forward.

II

By J. M. THYNE
(*Jordanhill College of Education, Glasgow.*)

The Scottish colleges of education were not involved in the ATCDE/BPS inquiry, but the Report is, of course, of professional interest to them, and to their psychology departments in particular. I comment, however, as the representative of only one, albeit the largest, of them—with a present full-time student population of over 2,500.

With the first part of the Report, most of my colleagues in the other colleges would be, I think, in general agreement. Child development (social, emotional, physical, intellectual, moral); the nature and conditions of learning as they relate to habits, skills, remembering, understanding, and transfer; the psychology of the basic school subjects; intelligence and its measurement, and applications to educational issues; these have long been major parts of the psychology syllabus for women on the three-year course. (Our departments of psychology are concerned also with graduate students, with students holding diplomas in art, music, domestic science and other technical subjects; but women on the three-year course form the largest group of students—over 1,600 at Jordanhill, this year—and it is to them that my comments here refer.)

What to include in the psychology course is, nevertheless, a major problem, not because it is difficult to find material relevant to the training of teachers, for it is not, but because of the difficulty of deciding what must be left out. The relevant material has to be arranged in a hierarchy of relevance, and a line has to be drawn in accord with the time available, the claims of other subjects, the abilities of the students, their own psychological maturity, and other things besides. In contrast with the Report, the line I draw excludes the training of students to apply tests, but it does include an attempt to cultivate in the student an understanding of the history, structure, and function of tests, and an ability to interpret in educational terms the results of tests which will be applied to their pupils by persons with greater expertise. This example raises the question of the criteria to be applied in selecting psychological topics, a point raised again in my concluding paragraph.

I am in sympathy with the caution expressed in the same section of the Report, though I have some reservations. As the Report says, argument by analogy from animal learning, and from highly controlled experiments on human learning, is of course, dangerous. But I, and many of my colleagues, would regard a deliberate attempt to minimise this danger as one of the fundamental tasks of the psychologist who undertakes to teach his subject to teachers. It is his special responsibility (not necessarily shared by his fellow-psychologists in universities) to ensure that what he says about learning (for example) is, in fact, valid for the kind of situation which will confront his students. And he will try to do this, not by ignoring the material in question, but by restructuring it, and testing in the school situation the results of his restructuring it. For example, analysis of human habit-forming indicates that its conditions of occurrence (if not its content) are basically the same as those made explicit in classical and instrumental conditioning. Also, it is not difficult to show that certain psycho-analytic mechanisms, such as Projection, occur within the lecture-room and the classroom. Caution is not enough; it must lead to constructive thinking. Admittedly, the touchstone of practicality serves not only to give reality to what is taught but also to reveal the sterility of what one might otherwise have taught. It is also a humiliating revealer of one's own ignorance.

I am in complete agreement with the Report's emphasis on the need to distinguish between educational psychology and educational philosophy, but I have learned to be pessimistic about the possibility of making that distinction clear to all students. Nearly all the Scottish colleges have a department of psychology and a separate department of education (an arrangement of which I heartily approve), but this physical manifestation of segregation does not always make apparent to students what the difference in function is. To take but one rather blunt example : in many education seminars students are encouraged (very properly, I believe) to express their own opinions and develop their own judgements ; whereas in a psychology seminar they may have to learn that the opinion of a student, as contrasted with his knowledge, is of little or no value. What matters here are the facts, and the hypotheses validated by reference to them. They may have to learn also that the observation of two or three children does not provide an adequate basis for the formulation of general hypotheses about child development. (Such observation can, of course, be very useful for other purposes)

On the central part of the Report (the actual content and forms of psychology courses) it is difficult to comment Lists of topics, even when accompanied by tables of hours, do not answer the questions that really matter. What topics are expounded in detail and not merely mentioned in passing, or in a somewhat general way ? And just what details are emphasised ? How much of what appears on the syllabus, or is referred to in lectures, seminars, and discussions, is the student expected to know and remember ? Is the material up-to-date ? Even the analysis and classification of examination questions (which must have been a formidable task) reveals remarkably little about the above issues, though some of the general questions quoted in Appendix 4 seem to imply that some parts of the courses must themselves have contained little detail. I am tempted to suggest that more precise information might have been obtained from an analysis not only of the examination questions, but also of the students' answers Good psychologists could scarcely object to a procedure widely used within their own discipline.

Probably the major difference between the Scottish colleges and the colleges involved in the Report relates to the qualifications of the teachers of psychology. All lecturers in psychology in the Scottish colleges of education have an honours degree of which psychology is a main subject All are trained teachers, with experience of teaching. Some have worked in child guidance clinics. In such a context I am bound to comment that the recommendations of the Report in this connection are unduly mild, though I do appreciate many of the practical difficulties.

The problem of suitable textbooks is serious, though slowly improving. So many of the texts which are sound on psychology have little or no bearing on the work of the teacher, and so many of those which are realistic about the classroom situation contain very little psychology—as psychologists nowadays would use that word. My own experience suggests that the ' general ' textbook of educational psychology may have to dispapear from training colleges. When reasonably compact texts on specific topics (such as Memory, Learning, Child Development, Intelligence) are available, they are often much more useful than the thinly spread omnibus volume. The ' specific ' texts allow not only of relevant selection but also of specialisation.

I have not attempted to pass comments on all parts of the Report, for such an attempt would occupy more space than it would deserve I think that the type of investigation described in the Report is valuable, and the committee responsible for it merits our thanks. Now, without in any way detracting from the value of that Report, may I suggest that a further task might profitably be undertaken, namely to discover the major problems actually faced by lecturers in psychology, and the ways in which they have been tackled ? For example, what criteria are cited in selecting psychological material for a course ? And why ? Is there an optimum order of presentation ? How do we solve the problem of breadth *versus* depth ? But it would be part of such an investigation to discover what questions should be asked.

COMMENT

ON

THE RELATION OF NEUROTICISM AND EXTRAVERSION TO INTELLIGENCE AND EDUCATIONAL ATTAINMENT

A brief comment on the controversy between J. B. Biggs and R. Lynn.

By H. J. EYSENCK

In the controversy between Lynn and Biggs, the crucial part appears to be played by the exact status of 'anxiety.' Biggs assumes that 'anxiety' and 'neuroticism' are used interchangeably by me, but this is incorrect. Neuroticism is regarded as an inherited psycho-physical disposition, closely linked with the lability of the autonomic system, which governs a person's emotional reactivity, and may predispose him to the development of neurotic disorders under suitable circumstances. Anxiety is a conditioned fear reaction which is particularly characteristic of dysthymic neurotics, i.e., of persons who are high on the factor of neuroticism and also on the factor of introversion, which is significantly correlated with conditionability. Anxiety, therefore, is a 'mixed' concept, being related both to neuroticism and to introversion. (The position is rendered rather confused by Cattell's use of the terms 'neuroticism' and 'anxiety,' which is exactly the opposite of mine.)

My theory, which links introversion and conditionability, is not necessarily antagonistic to that of Spence, which links neuroticism, regarded as a drive, to conditionability. On the theoretical level either theory, or both, or neither, might be right. My belief that Spence's theory is inadequate is based on the experimental evidence, and may very well be mistaken; the experimentally demonstrated correlation between the Manifest Anxiety Scale and tests of introversion make all the Spence-Taylor experiments using the M.A.S. equivocal. In our own work we have never succeeded in finding correlations between neuroticism and conditionability, and neurotics as a group are more conditionable than normals *only* when they are selected in such a way that introverts are over-represented.

I agree with Biggs that a great deal more work is needed at the theoretical level, but I feel even more strongly that what is needed is a greater concentration on *experimental* studies, with particular reference to parametric investigations. The frequently contradictory results obtained by investigators testing both Spence's hypothesis and my own suggest that neither hypothesis is universally valid, but requires to be closely circumscribed with respect to its validity by detailed specification by such parameters as strength of CS and UCS, CS-UCS interval, proportion of reinforced trials and so on. Lynn's study appears to me to be a logical deduction from principles which appear to have at least some experimental support, although this support may be rather circumscribed in the manner just outlined, and although it may, of course, be possible to account for his results in other terms than those premised in his theory. This hardly detracts from the value of his important demonstration.

CRITICAL NOTICES

HUNT, J. McV. (1961). *Intelligence and Experience.* New York: Ronald Press, pp. vii+416, no price given.

BY P. E. VERNON.

This is an important book on at least two counts. First, Professor Hunt puts forward the strongest case yet made for discarding the conception of intelligence as inborn potential which is predetermined by the genes and which matures regardless of environmental conditions. Secondly, he provides the most complete exposition available in English of the whole corpus of Piaget's theories and experiments on mental development, which lead him to conclude that intelligence should be regarded as " central processes comprising strategies for processing information that develop in the course of the child's interactions with his environment."

He argues that predeterminism has been implicit in developmental psychology generally, for example in the conventional distinction between maturation and learning, and in the advice given to parents by mental hygienists not to ' push ' their children's intellectual growth. So strong was its influence till recently that the common reaction to experimental demonstrations of the modifiability of measured intelligence (e.g., those of the Iowa school in the 1930's) was one of derision and emotional controversy. But with the appearance of Hebb's work on the build-up of autonomous central processes and the effects of varied environmental stimulation, Piaget's descriptions of the mechanisms of accommodation and assimilation in the progressive developments of schemata, Harlow's experiments on the acquisition of learning sets, and many other lines of evidence, the whole theoretical climate has changed. Besides describing this work in detail, the book is quite fair in recognizing contrary evidence, such as Gesell's classical co-twin control experiments on maturation, and shows how these can be reconciled with the newer theories.

Hunt does not deny that genetic factors set a limit to the complexity of what can be learned, but his interpretation of the work of modern geneticists leads him to deny that we can ever isolate and measure the genetic component either in physical or mental development. Attempts to apportion hereditary *versus* environmental variance (e.g., the often quoted 80 per cent.—20 per cent.) are misleading, since at best they only reflect the environmental pressures that have operated in the sample being studied, and tell us nothing about what would happen under different environmental conditions. Many investigations of more extreme environmental deprivations indicate far greater effects, and he goes so far as to predict that the application of a sound psychology of development could raise the average intelligence by an amount equivalent to 30 points of I.Q. In the reviewer's opinion, this forecast is unnecessarily rash; we know far more about the conditions that inhibit than those that favour the full realization of intellectual potential. And, although Piaget's work has thrown a flood of light on cognitive processes, it does little or nothing to improve our control of the even more important motivational ones.

The main key to such improvement, he claims, is the matching of environment to the current level of the child's schemata; (he admits that this principle is already realized by the good school teacher). If the child is set problems which merely exercise schemata that are already organized, he gets bored; while, if they are too far ahead of him, he is uninterested or frustrated; whereas with appropriate choice of stimuli, there should be positive pleasure in learning just as we find in the baby's acquisition of talking and walking. Circumstances which are appropriate at one stage may inhibit at another; for example, the highly protective mother may provide adequate experience in the sensorimotor stage, yet deprive her child of it in the preconceptual (1½–4 year) stage. Variation of circumstances is another important principle in the development of flexible and transferable skills.

These educational implications are among the most original parts of the book. For the most part it covers ground which is already fairly familiar to educational psychologists in Britain (though perhaps less widely known in America). As is only to be expected from a polemical treatise, the author's enthusiasm sometimes runs away with him. For example, it would have been helpful if he had given some consideration to the strong counter arguments that have been put forward by Professor Burt in a series of recent articles in the General and Statistical Journals of the B.Ps.S. Again, he probably gives too much credence to Piaget and Inhelder's claim that each stage develops out of preceding ones (e.g., that realization of constancy of quantity always precedes constancy of weight, which always precedes constancy of volume); and that a wide range of restructurings tend to occur simultaneously (e.g., that the attainment of formal thinking about scientific problems is accompanied by the revolutionary attitudes about society characteristic of adolescence). The more carefully controlled studies of Peel and Lovell, for example, indicate much more variation and inconsistency than Piaget usually admits. However, there can be no doubt that this will be in influential source book for many years, particularly since, although fairly technical an parts, it is a good deal easier to follow than Hebb's or Piaget's own writings. It is strongly recommended to advanced students and to lecturers in education and psychology.

THE APPRECIATION OF BEAUTY

VALENTINE, C. W. (1962). *The Experimental Psychology of Beauty.* London: Methuen, pp. xiv+438, 45s. net.

BY CYRIL BURT

Professor Valentine's aim has been to present a systematic survey of experiments on the appreciation of art, music, and poetry, and to determine how far the psychological differences between individuals helps to account for the wide differences in what is known as 'artistic taste.' Those who have acted as subjects in these various researches have frequently been children of school age. Hence, quite apart from the intrinsic interest of the whole problem, the observations recorded and the conclusions drawn have a direct bearing on the work of teachers.

He begins by observing that philosophers, art critics, and aestheticians have so far failed to agree on what is meant by the word 'beauty', and he himself declines to offer any formal definition. Many, indeed, would hold with G. E Moore that, strictly speaking, the word 'beauty,' like the word 'good,' is indefinable: the most we can do is to recognize it when we see it, and point to typical instances which may serve to indicate what precisely we have in mind. In ordinary life nearly all of us distinguish certain conscious experiences as having a common quality which we indicate by the term 'beautiful,' although in other respects they may vary widely: they may be visual, auditory, verbal, emotional, or even purely intellectual. Beauty itself seems a special form or aspect of what is technically termed 'value'—a word which is interpreted either objectively (as the older writers usually did) by saying that it denotes "something that is worthy to exist," or subjectively, by saying it is "what we prefer, like, or admire." But once again, these are merely rough, ostensive descriptions. Value itself is a 'primitive concept,' and as such cannot be defined. Yet that need not prevent scientific study. Vague as they are, the descriptions given are sufficient to show what sort of experiences we propose to study, what types of material we must therefore select, and the kind of questions we need to ask.

The characterization to which, it would seem, the experiments themselves eventually lead (so Professor Valentine tells us in his final chapter) is that the experience of beauty is distinguished by a "delighted absorption in the object itself." One thing becomes increasingly clear. "Art," as Herbert Read has emphasized, "is not necessarily beauty · that cannot be said too often or too blatantly." Nor is there any unique type of feeling that we can label *the* aesthetic emotion. The experience seems rather to be a perception or perceptual judgment of some particular quality, or combination of qualities, which is as a rule accompanied by " a harmonious blend of feelings and a certain attitude of mind." The unexpected complexity of the whole process and the astonishing number of factors that affect it are perhaps the most instructive conclusions of the entire series of investigations that Professor Valentine describes.

His book is divided into three parts—the first dealing with visual experiences, the second with auditory, and the third with verbal. In each he begins with the simplest types of stimuli, and advances progressively to the more complicated. Accordingly, after a brief introduction, his first chapter starts with experiments on simple colours, and reports the results of numerous studies carried out by himself and others on colour-preferences among infants, school children, and adults from various races and countries. So far as comparable data can be obtained, it appears that during early childhood vivid colours are everywhere strongly attractive, and that red, orange, and yellow are generally preferred to green, blue, or violet. Later, green or blue tend to predominate, and white rises high in order of preference. Here at the very outset we encounter marked individual differences, and detect the operation of a variety of factors, similar to those we meet with in experiments on more complex objects. Thus, in one set of investigations it was found that children as they grow older become more and more influenced by personal association, particularly in the case of colours that are unsaturated or ' mixed ' rather than saturated and ' pure ' : one boy, for example, says he likes a pale faded yellow-green, " because it makes him think of leaves in Autumn " ; another dislikes red, " because it reminds him of blood." These illustrate the commonest or ' associative type.' Other children refer to the ' subjective ' feelings the sight of colours arouse : " this yellow dazzles the eye," " that bright red makes you feel hot all over," " the green is so restful and soothing." Somewhat rarer is the so called ' character type ' : one girl, for instance, speaks of scarlet as " loud and aggressive " ; another of pink as " rather demure." And finally, a few among the oldest adopt an entirely ' objective ' attitude, and give reasons relating solely to the colours themselves : they like an ultramarine "because it is such a pure blue," and dislike a certain green "because it has too much yellow in it to be a really good green." As Professor Valentine points out, these four types, with varying frequencies, crop up in every medium and at almost every stage. Nevertheless, so he maintains, there is sufficient uniformity to suggest, in addition to the effects of experience or temperament, the influence of an innate differential response to colours of different hues, changing spontaneously as the child matures. From single colours we turn to colours in combination ; and at once we find that the mere presence of another hue can greatly modify the initial preferences. However, the appreciation of colour-harmony and the distaste for colour-clash does not appear until relatively late—six or seven years, according to most observers, though with a few individuals it emerges as early as four.

From colour we proceed to form. There is an unexpected measure of agreement over preferences for certain kinds of lines, curves, and linear shapes. Here the influence of *Einfühlung* or ' empathy ' is particularly striking : the spectator, so to speak, sympathetically puts himself in the place of the lines and shapes ; he feels himself moving along the graceful curve or supporting the heavy horizontal oblong. ' Subjective types ' and ' character types ' are thus strongly in evidence. Certain zigzags are described as ' lively ' or ' playful,', others as ' harsh ' or ' cruel.' Rising lines seem ' active ' or ' hopeful ' ; falling lines ' weary ' or ' sad.' And the experiments sufficiently prove, if proof were needed, that there is no single line or curve which is unanimously preferred. The recurrent search by theorists like Lomazzo, Hogarth, and Winckelmann for a so-called 'line of beauty' ended, we may remember, in the serpentine curve of a question-mark.

From the canons of Polyclitus and Vitruvius onward to the present day, numerous writers have maintained that, as in music and poetry, so in art and architecture, the secret of beauty lies in the component proportions, and that these proportions could be numerically expressed. The doctrine is most succinctly summed up in a dictum of Palladio (*I Quattro Libri dell' Architettura*, 1570) : " the beauty of a structure consists in the exact proportions within the parts themselves and of each part within the whole." It is, therefore, not surprising that the earliest experiments on the appreciation of the beautiful—those carried out nearly a century ago by Fechner, " the father of both psychophysics and experimental aesthetics "—were concerned primarly with the problem of the most pleasing proportion. He began with the simplest cases—the cross and the rectangle.

Many teachers will remember puzzling over Definition 3 in Book IV of Euclid's *Elements* : " A straight line is said to be cut in extreme and mean ratio when the whole is to the greater segment as the greater segment is to the less " : i.e., if a and b are the two segments, $a/b = b/(a+b)$; (the solution is $b = (\sqrt{5} \pm 1)a/2$; for the construction, see Euclid VI, 30 and II, 6). This ratio defines what artists of later age called the ' Golden Section.' The proportions of a common filing card, 3×5-in., are a rough approximation ; 5/(3+5) slightly nearer ; and 5/(8+5), the dimensions of a foolscap sheet of typewriting paper, nearer still. This sequence—1/2, 2/3, 3/5, 5/8, 8/13, . . . , (based on the so-called ' Fibonacci numbers ') progressively approximates the golden section—the proportion adopted again and again, sometimes consciously, sometimes unconsciously, by practising artists, architects, and sculptors. The rival theory maintained that the ratios for visible beauty should be the same as for audible beauty, namely, those which determined the most harmonious intervals in music—1/2, 2/3, 3/4, 5/6 and 3/5. Both theories held that the simpler ratios served for the simpler structures, and that the more complex could be appreciated only by the relatively sophisticated spectator or listener. Fechner claimed that his experimental results showed plainly that the formula for the golden section furnished by far the most pleasing proportion, and gave ' the right combination of diversity in unity."

Professor Valentine reviews the numerous researches since carried out with subjects of different ages to verify Fechner's conclusions At first sight the data obtained seem hopelessly at variance. Nevertheless, the main trend is roughly in keeping with a generalization suggested by one recent investigator, Miss Donaldson. She found that up to the age of about 5, children exhibit no definite preference ; then they tend to favour the simplest ratio of all—1/2—i.e., an oblong which is a double square ; later, they prefer a ratio approaching 2/3 ; later still 3/5 ; and so on, until at about the age of 12 the golden section becomes the most popular (except when the oblong suggests some specific object such as a window or door) ; after that, even this rather subtle relation begins to appear flat and insipid, and more daring proportions, varying widely with different individuals, come to be preferred, with the broader ratios (3/4 or 4/5) predominating. This tallies with the verdict of Bacon. " There is," he writes, " no excellent beauty that hath not some *strangeness* in the proportions : a painter may make a better face than ever was, but he must do it by a kind of felicity, and not by rule." And certainly one can hardly suppose that any strict mathematical sequence governs the development of the child's preferences or that any exact numerical ratio determines our maturer appreciation of beauty.

R. M. Ogden, it is true, professes to find golden sections, pentagons, ellipses, and circles, nearly always underlying the idealized human figure as represented in the paintings of Renaissance artists and the statues of ancient Greece. But, as Professor Valentine rightly comments, if you draw a sufficient number of geometrical shapes over your picture, some of them are bound to provide an approximate fit. Miss Puffer's experimental results, which he discusses in greater detail, appear to come nearer the truth. The proportions involved are not just geometrical ratios in space : they express, as it were, a balance of forces. Instead of the simple symmetry that appeals to the younger child as he arranges objects on the mantelpiece, the experienced eye prefers " that subtler type of composition, which symbolizes some sort of equilibrium "—particularly, I would suggest, the homoeostatic principle so characteristic of the *living form*.

This, I take it, is why the nude has played such a vital part in art and art-training. The ' Venus coelestis ' of the Greeks (the Cnidian Aphrodite, for example), the ' Venus naturalis ' of the Gothic convention (as exemplified in the ' Eves ' of Memlinc or Van Eyck), the sculptured pharaohs of the 18th Dynasty, the David of Michael Angelo, and the beardless and undraped Christ who pronounces sentence in the Last Judgement—all are based on proportional schemes; and yet their several proportions differ as widely from each other as they do from those of the male or female model in the life-class. Hardly any one in my experience, whether educated, illiterate, or avant-garde, ranks even the most artistic photographs one can find either of naked athletes or the winners of the latest beauty contest as more pleasing aesthetically than the relatively unrealistic figures of antiquity or the high renaissance; and, what is still more significant, even when the inexpert spectator fails to distinguish the photograph of a statue or painting from the photograph of a living model, the former occasions none of the embarrassment so often caused by the latter. Each work of art is, as we say, 'idealized'—in the sense that the shapes and planes, the anatomy and the general poise, are so simplified as to help the spectator to apprehend in a glance something of the tense play of forces that underlies the most complex and highly developed forms of life. For the picture or the statue itself must have something of the characteristics of a living and organic whole.

In the whole field of experimental aesthetics by far the most popular subject for research has been the appreciation of pictures. One result, on which all investigators are agreed, is the way the content of the picture—the story that it tells or the subject which it depicts—governs the preferences of younger children and less cultured adults For the rest one's first impression is that so many different factors influence each person's artistic judgment, even among the expert and the connoisseur, that any attempt to extract some underlying principle is completely swamped by the overwhelming fact of individual differences. *De gustibus non disputandum*. However, when a systematic statistical analysis is undertaken, the proverbial maxim quickly proves to be grossly misleading. If we collect a set of (say) fifty pictures ranging from good reproductions of acknowledged masterpieces, second and third-rate specimens of every type and variety, down to the crudest birthday card to be found in the local paper-shop, and then compare the order of preferences obtained from a representative sample of people, including both the cultured and the uncultured, the philistine and 'the fellow from the bus,' as well as the professional artist and art critic, we find that almost all the correlations between their rankings are positive, and the average coefficient surprisingly high To borrow the technical jargon of the mental tester, there is a well marked 'general factor' accounting for about 50 per cent of the 'variance.' That, of course, still leaves ample room for wide individual differences. Subordinate factors indicate the presence of supplementary tendencies or types—some persons, for example, being guided mainly by colour, others by form; and to a large extent, their type of preference tends to be correlated, not only with their past experience and training, but also with their temperamental and emotional outlook.

Professor Valentine concludes this section with a brief but instructive chapter recording experiments carried out by himself and his co-workers on 'reactions to modern art.' Here, the individual differences were even more conspicuous. In general, among the 'unsophisticated students,' who formed the bulk of his subjects, conservative or conventional judgments were commonest. Nevertheless, an unexpectedly large number described the pictures as 'interesting,' either (so it would seem) because of their novelty or their daring or because of the bewildered curiosity they so often excite. Those who like them most are often high-spirited youngsters who find the stock types of beauty dull, effeminate, or devoid of interest, and enjoy having their tougher emotions stimulated and admire those who have the courage to violate the traditional rules "At times," as Eric Newton observes, "it is refreshing to smash the china and break bottles, or watch others at the game." Occasionally, one even wonders whether the modernist does not deliberately play up to this attitude: indeed, Professor Valentine himself quotes a confession attributed to Picasso to the effect that he had, in his later years, intentionally "fed those fellows with what they wanted, and become rich and celebrated through amusing myself with these farces."

The appreciation of music is dealt with along similar lines, but rather more briefly. The experiments begin not with single notes (e.g., those played on instruments of different timbre), but with simple combinations of notes, i.e., with a comparison of intervals. Professor Valentine's own experiments on this and kindred problems—first with school children and then with students—were started more than fifty years ago. He found much the same types of attitude here as in dealing with colours. Moreover, contrary to the assumptions commonly made, the consonant intervals were by no means always the most pleasing, nor the dissonant the most unpleasant. At the present day, the average order for most persons appears to be—major 3rd, minor 3rd, major 6th, minor 6th, octave, perfect 4th, perfect 5th, augmented 4th (tritone), major 2nd, minor 7th, major 7th, minor 2nd. This, it will be seen, corresponds neither with the order suggested by the ratios nor with that suggested by the relative absence of beats. The ages at which children can distinguish concords and discords, as well as their actual preferences, are greatly influenced by early musical training; and during the course of a single set of experiments repetition and familiarity will often render a discord as agreeable as the concords.

The long chapter on 'special factors in listening to music' is one of the most interesting in the book. The chief problems discussed—always in the light of experimental evidence—are the influence of specific characteristics, like rhythm, melody, harmony, and timbre, on the listener's general appreciation, and the effect of the general characteristics of the music on his more specific responses—on his imagery, for example, and on physical processes such as pulse, respiration, and (most obvious of all) definite movement—nodding, beating time, dancing, or marching more rapidly or more slowly. The marked effects on speed of movement suggests that an instructive research might be carried out on the converse influence—that of the natural motor tendencies of a listener, composer, or conductor on the style of performance he favours Few conductors, for example, keep strictly to Beethoven's metronome markings; and their differences in regard to tempi are amazing. Beecham, at the Festival Hall in 1956, took barely an hour to get through *The Ninth Symphony*; Hindemith, a couple of years later, took nearly an hour-and-a-quarter Now it has been claimed by more than one musical critic that the preferred rates of both composers and conductors are noticeably influenced by the length of their limbs; the tempi of Beecham and Kleiber, for instance, are on an average much faster than those of Toscanini, Furtwängler, or Klemperer; and the numerous allegros of Bach and Berlioz, as contrasted with the ambling andantes favoured by Schubert, have been associated with the speed at which, according to their biographers, each composer habitually walked. Analogous differences are sometimes discernible with tall and short listeners, though I know of no recorded correlations. This is merely an incidental example of the numerous out-of-the-way factors that keep cropping up as soon as one starts watching for possible influences. May we hope that music teachers will not only study the suggestive conclusions that Professor Valentine himself has drawn, but will also exploit their exceptional opportunities for carrying out similar researches of their own?

The concluding chapter in this section is concerned with music as a " vehicle for expressing meaning and emotion " Many composers, particularly the romantics, have openly declared that their aim was to convey emotions, represent scenes, or even portray characters (as in Strauss's *Till Eulenspiegel* or Elgar's *Enigma Variations*). But the outcome of numerous experimental researches is to demonstrate how hopelessly even the most expert composers will often fail. And, as Professor Valentine points out, a number of investigators have accordingly concluded that music neither can nor should seek to communicate a visual, verbal, or emotional 'programme'; many musicians would, of course, agree. But the criteria adopted by the investigators are, perhaps, too severe. The composer would probably not claim that his audience should be able to guess his intention with no further clue, but merely that the style of his composition should be *consistent* with the mood he attempts to convey, which he commonly indicates by a title. Moreover, he would usually expect that the competent listener would already be familiar with the conventions and devices which he and his predecessors had developed for particular purposes.

Doubtless, there are many things that music cannot effectively communicate. But numerous results cited by Professor Valentine show that, in many cases, certain emotions and moods are expressed quite successfully. For example, in one investigation 93 per cent. of the subjects, given a list of alternative descriptions, agreed that the ' Marcia Funebre ' of the *Eroica* expressed grief, sadness, or despair (only 12 per cent. could identify it) ; 63 per cent. gave triumph or determination as the mood expressed by the last movement of the *C Minor* ; 72 per cent. agreed that Nicolai's *Merry Wives of Windsor* expressed amusement or joy, and so on

When I myself have quoted these or similar examples in lectures or discussions, I am nearly always asked ' Why ? ' So may I briefly mention the answer I should give. The human voice is our supreme medium for expressing emotion ; and in so doing it varies recognizably in level and change of pitch, in level and change of loudness, in level and change of tempo, and in changes of timbre. Now these are all distinctive qualities of music. Where no special emotion is expressed, the melody at its simplest tends to follow the stock pattern of speech used in making plain affirmations, i.e., its basic ' intonation ' as the phonetician calls it : each ' sentence ' rises and then falls, and the sentences themselves rise to a climax and then drop to the original level ; emphasis can be added by special stress. The most familiar instances are to be found in well known hymn tunes—e g., *Tallis* (' The heavenly child in stature grows '), *Innocents* (' Gentle Jesus '), or *Petra* (' Rock of Ages ') ; but it is also found, with added emphasis, in popular songs (e.g., ' We won't go home till morning ') and in classical melodies (e.g., ' Freude, schöner Götterfunken,' in the *Choral Symphony*). It is likewise the scheme of the old psalmodic plainsong and of ordinary recitative (*musica parlante*). If a command, a direct question, or a question with an interrogative is to be expressed, the melody tends to follow the rise or fall that would be used in ordinary conversation. Gaiety and excitement are indicated by a raised pitch and a quickened tempo ; rest or sadness by a lowered pitch and a slow tempo. Vigour is expressed by wide intervals (e.g , the drops of a seventh in the ' Nimrod ' variation) ; weariness or anxiety by the smallest possible intervals (e.g , the opening phrases in the overture to *Tristan*) ; and note that in both cases a suitable difference in timbre is produced by the choice of instruments and the unusual harmonization. But I need not multiply the parallels.

Of course, in highly developed music these echoes of speech-melody get overlaid by the distinctive characteristics of the musical tradition to which the composition belongs—the devices of key, scale, melody, rhythm, modulation, harmonization, and the formal structure, which the composer inherits or chooses to introduce. Music as such is an art with principles of its own ; and most musicians, I imagine, would hold that, even if an incidental or concomitant expression of definite emotions may at times form part of its aim or technique, this is not its primary purpose. Its function is to produce aesthetic compositions (in the literal sense of the word ' composition '), patterns put together out of the relevant material—in this instance sounds with their varying qualities of pitch, loudness, timbre, etc., ***and their emotional associations.*** If, like other works of art, it arouses a specific feeling, the essential feeling (as the listener's comments sufficiently prove) is aesthetic feeling.

Professor Valentine devotes only a few pages to " responses to modern music." For some unexplained reason, modern music, unlike modern painting or modern poetry, has failed almost entirely to stir the interest of experimental psychologists. With one exception, which he mentions in a footnote, he believes his own experiments are almost the only attempts in this novel field His chief results are, first that the judgments passed by the listeners are based most frequently on an ' objective attitude ' and comparatively seldom on ' associations '—quite different from what is found in the case of traditional or classical music, and secondly that, in spite of the marked differences " there are far more similarities in the responses " than we might have expected. From my own experiments it would seem that young people with little musical experience and no musical training rarely appreciate modern music *as* music. Many describe the irregular and the primitive character of the rhythm as ' exciting ' or ' stimulating ' , a few feel the strange harmonies or lack of harmony to be productive of a heightened tension : one speaks of it as ' tormented and rather neurotic ' ; several, particuarly among the younger and more rebellious males, find it exhilarating to get away from what they term the ' sloppy ' or ' namby-pamby sentimentality ' of the older music—" all pretty-pretty and rather stale," as one of my listeners observed. Most of those who liked it were already vaguely familiar with modernism from signature-tunes and accompaniments in the cinema or television programmes. The rest thought it ' difficult to follow ' and ' tuneless.' Of those who had enjoyed some musical training, a number described the compositions as ' interesting experiments,' and many said they would understand and enjoy them better after further hearings. The more appreciative fell into two main types : the larger group preferred the so-called neo-classicists (as represented, for instance, by the later works of Stravinsky) ; the smaller group preferred the atonalists (Schönberg and his various followers), though none used these terms or identified the composers.

However, what the psychologist wants to know is this. Not one of the modernists, I imagine, would accept the common type of criticism—that he is merely experimenting, merely trying to be original or to show what ingenuity can accomplish with a twelve-tone scale, or even just striving to escape from the tedium of the old classical rules and the " linked sweetness long drawn out " of the later romantics. Nearly all would contend that, with their revolutionary principles—rejecting the stylized structure and the call for a clear melodic outline, the traditional formulae for resolving discords, and the adherence to definite keys and definite time-signatures—they are able to achieve very much more than any of their hide-bound predecessors. But, we must ask, have they succeeded in producing the effects they intended ? Or are their intentions comprehensible and perceptible solely to those esoteric followers who have expressly studied the new fashions ? On the whole, I am inclined to think that the experimental results undoubtedly show that they have, in some measure, been genuinely successful. And yet, I would suggest that further investigations of the kind Professor Valentine describes are really needed before a psychologist can feel fully convinced. One result, I suspect, would quickly emerge : the approach adopted by nearly every theorist, from Rameau and Monteverdi to Helmholtz, Prout, and Schönberg—all of whom base their doctrines primarily on mathematical and physical considerations rather than on psychological—is both mistaken and misleading.

And so we come to the final section of the book, the shortest of all, which deals with the appreciation of poetry. Once again I find Professor Valentine's own experiments by far the most interesting of those that he records. And where he disagrees with the conclusions of the literary critics—particularly those of the well-known studies of I. A. Richards at Cambridge—Valentine's views always seem more convincing than theirs. In his work with children he finds (as we might expect) that at the younger ages, " the story is the main thing," and yet (what we might not have guessed) many reveal " a surprising sensitivity to good and bad poetic expression " : " the sense of beauty develops even before the power of clear understanding." During earlier years rhythm exercises the most potent influence ; the appreciation of word-melody and vowel-music comes much later. Visual imagery may often enhance enjoyment, though occasionally it impairs it. He gives several instructive instances to illustrate the child's approach. Wordsworth, for example, in the first edition of his ' Sketches among the Alps ' has the line : " Under a hoary oak's thin canopy ." Later he changed it to " Beneath an old grey oak as violets lie." Young children prefer the former version, " partly because, as one of them puts it, ' hoary is a nice poetic word ' " ; the older usually prefer the second, " because it makes a pretty picture, and sounds best ; the ' hoary ' oak and the ' canopy ' seem like using grand words for effect."

Professor Valentine adds a further chapter on ' experiments with modern poetry.' His own appear to be the only investigations on the subject carried out in this country, though a few have been attempted in America. Most of his students commented on the difficulties presented by such poems, and there seems to be a correlation of 0·2 and 0·3 between the ability to comprehend them and the student's I.Q. However, T. S. Eliot, we may remember, observes that " the seasoned reader does not bother about understanding—at least not at first." Contrary to common opinion, the experimental results show plainly that " it is quite possible for people whose favourite poets are Wordsworth, Keats, and Pope to find some modern poems also attractive." But the main conclusion that Professor Valentine seeks to emphasize is that here, perhaps more clearly than anywhere else, we find that " art is not necessarily to be identified with beauty." The aim of such poetry is to express, not so much a sense of what is ordinarily considered beautiful, but rather an " intellectual toughness in the midst of passionate feeling " ; and he quotes with approval the remarks of Goethe that " no object need be excluded as unpoetical" and of Daiches that the merit of the modern style is " to render, in a way no other medium can achieve, a particular apprehension of the excitement and complexity of experience."

The book ends with a chapter of conclusions, recapitulating the main results and discussing what general inferences are common to all three fields. The first and most obvious conclusion, which emerges from almost every experimental study here reported, is the wide variety and range of the individual differences. And yet, underlying these differences, certain basic principles appear, especially when the data are analysed statistically. To begin with, in each of the three fields, there are large general factors commonly covering all the persons tested . thus, in spite of their superficial peculiarities, each person tends, in a greater or lesser degree, to rank the items presented in much the same order of merit. Next, we find another general factor, almost as large, covering all the different fields : thus, on the whole and in the long run, a person who is a good judge in one field tends to be a good judge in the others. This capacity is partly, but by no means entirely, accounted for by the general factor of ' intelligence.' There seems also to be a general capacity for aesthetic appreciation as such. But, in addition to this general capacity, there are several specific factors, temperamental (so it would appear) rather than cognitive or intellectual—for example, those giving rise to a preference for romantic rather than classical works, and again for realistic rather than impressionistic techniques ; and each of these seems to influence appreciation in all three fields, visual, audible, and verbal. Moreover, the rule that " attention should be at once stimulated and facilitated " is common to all three types of artistic creation. Thus, in painting, in music, and in poetry the principles of composition are very largely the principles of attention : there must be sufficient variety, complexity, and novelty to excite and retain the observer's attention ; and at the same time, there must be sufficient unity and familiarity to prevent his attention being completely baffled. The span of the observer's attention—a span which, as we know, increases with the increase in the child's mental age—must itself be a factor determining and limiting the observer's appreciation. And, of course, the capacities for discrimination in different sensory fields—visual, auditory, and verbal—are also relatively specific, developing or maturing at rather different ages. Each of the various capacities and temperamental characteristics revealed by these experimental results must be to a large degree innate. But how much is due to genetic constitution and how much to past experience and to cultural environment is a problem on which there have been many dogmatic assertions, but, up to the present, little or no scientific investigation.

It is now just half a century since Professor Valentine published a slim manual with the same title as his latest volume Since then, strange to say, no other book has been issued dealing with the same general topic. Professor Valentine's thus remains unique, and is certain to take its place at once as a classic. It is a book which will not only have a strong appeal for the interested layman ; it is one which every psychologist and every teacher should study. The increase in its size amply testifies to the vast amount of work that has been accomplished during the intervening years. Yet there are numerous fields which the experimentalist has scarcely touched—the appreciation of prose, of architecture, of industrial and domestic design, of the aesthetic aspects of the theatre, the film, and the presentation of television programmes Thus the urgent need is still for further research. This, as Professor Valentine rightly insists, is a task in which the artistic expert, critic and creator alike, should co-operate with the trained experimentalist far more closely than hitherto. And no one, I would add, is so well situated for undertaking and assisting in such inquiries as the teacher in the classroom or the lecturer in the training college or the schools of art and music.

BOOK REVIEWS

AINSWORTH, MARY D., *et al* (1962). *Deprivation of Maternal Care.* Geneva: World Health Organisation, available through H.M.S.O., pp. 165, 12s.

This book "attempts a reassessment of the effects of maternal deprivation in the light of advances in the last ten years and criticisms of the theory." It consists of contributions from the following: D. G. Prugh and R. G. Harlow—*Masked Deprivation in Infants and Young Children*; R. G. Andry—*Paternal and Maternal Roles and Delinquency*; M. Mead—*A Cultural Anthropologist's Approach to Maternal Deprivation*; B. Wootton—*A Social Scientist's Approach to Maternal Deprivation*; S. Lebovici—*The Concept of Maternal Deprivation—A Review of Research*; and M. D. Ainsworth—*The Effects of Maternal Deprivation—A Review of Findings and Controversy in the Context of Research Strategy.*

Inevitably in a book of this nature there is a certain amount of duplication; this is particularly evident in the reviews of research. The contributors have been chosen from widely different fields, however, and it is interesting to see differing interpretations of the same researches. Some of the contributions seem less significant and useful than others. Lady Wooton's, for example, is, in her own words "in the main an abbreviated report of what I have said already in the relevant chapter of 'Social Science and Social Pathology,' and adds little. Andry's contribution is useful in that it draws attention to the paternal role—as well as the maternal—but a purist might quibble about a contribution to 'Deprivation of Maternal Care,' which devotes at least half of its length to a study of 'the pathological role of fathers in delinquency'."

As a whole, this book is useful in crystallising and clarifying the many different problems generally subsumed under the heading 'maternal deprivation.' A number of the contributors rightly insist on the importance of distinguishing the different sets of conditions which are commonly and loosely discussed in terms of maternal deprivation. Dr. Ainsworth, for example, distinguishes—*insufficiency* of interaction implicit in deprivation, *distortion* in the character of the interaction without respect to its quantity, and the *discontinuity* of relations brought about through separation.

Dr. Ainsworth's article is, indeed, probably the most important and useful of the contributions; certainly it is the longest, occupying more than a third of the length of the book. Its excellence lies first in its detailed discussion of the separate questions arising from the concept of maternal deprivation; these include the questions of multiple mothers, variability in the degree of damage following deprivation, specific *versus* general effects of deprivation, diversity in the nature of the adverse effects of deprivation, permanence of the effects of deprivation delinquency and maternal *versus* environmental deprivation; secondly, in her very full review of the various 'research strategies' which have tried to answer these questions. This review includes comments on the preceding contributions.

Her article, as indeed do all the others—includes a very full bibliography.

In summary, a useful book drawing together, as it does, a mass of research evidence, presented and interpreted by authorities with widely differing *theoretical* backgrounds.

D. EVANS.

APLEY, J., and MACKEITH, R. (1962). *The Child and his Symptoms—A Psychosomatic Approach.* Blackwell Scientific Publications, pp. 262, 25s.

It is not normally the practice of this *Journal* to review medical publications, but a paediatrician colleague suggested that this book would be of considerable interest to psychologists who work with children. The book is by two distinguished paediatricians and is meant mainly for doctors. It is not a text-book of child psychiatry, but rather a comprehensive account of childhood disorders which, it is rightly claimed, cannot be understood or treated unless the physical aspects of the condition are considered alongside the intellectual, emotional and social. The claim is not, of course, new and all doctors would claim that they do consider their patients as whole persons, but they do it by the light of nature and the authors do not think this is enough. Rather they point to the need for doctors, particularly those who deal with children, to *study* scientific psychology, for, they point out, " the science of how humans behave is not just an attitude of mind, but an available collection of knowledge and techniques." The importance of correct and complete diagnosis is stressed throughout, with equal emphasis on the physical and psychological aspects, for while " the primarily physical disorders of childhood are often transient . . . for psychogenic disorders the prognosis must be more guarded."

Many, if not most, of the conditions dealt with in this book are those which are frequently found among children who attend child guidance clinics. They include among others, tics, eneuresis, recurrent fevers, cyclical vomiting, recurrent headache, stomach ache and other pains and the various problems associated with feeding. Often they are dealt with jointly by doctor and psychologist, sometimes by one or the other, sometimes a psychiatrist or other psycho-therapist is involved. This little book should do quite a lot to provide a common meeting ground. It will obviously help the family doctor to see the need for studying his patients more closely and to be on the look out for psychological concomitances and courses. It will help the child psychologist to look beyond his discipline and be ready to seek help of a paediatrician

The book is written very simply and clearly, case histories and quotations from the literature are sufficient in number and length to illuminate the text without ever becoming tedious. The authors are very modest and never dogmatic, they are not afraid to admit they do not know. They write, for example, " Because of our lack of experience in treating warts we offer no advice." They are also fully aware of what is normal in the physical, social and emotional development of children so that they are well able to offer caution about those conditions which children may or may not ' grow out of.'

There is no doubt that many readers of this *Journal* will find this a very valuable book both for general reading and for reference to special syndromes specific to children under their care. L. B. BIRCH.

GINSBERG, M. (1962). *On the Diversity of Morals.* London: Heineman, pp. xiv+329, 12s. 6d.

This book is a collection of papers originally published between 1932 and 1955. They form what is primarily a study in the philosophy of moral discourse and of sociology; and they cluster round certain themes. Ginsberg is concerned to examine what he calls " the possibility of a rational ethic." He regards the distinction between questions of fact and of value as valid and important, and he rejects a naturalist interpretation of ethics. Naturally, therefore, one of the chief problems that faces him as a sociologist is how to account for the diversity of morals manifested by different communities, present and past. This diversity has served in an important way to support the doctrine of ethical relativity (namely, " that moral value is relative to the subject asserting it or to the type of society in which it is commonly asserted," with the implication " that there is no rational way of deciding between differing moral judgments and, indeed, that the distinction between true and false does not properly apply to them.") Ginsberg is much concerned to show that the facts of moral variation do not consititute a good, let alone a conclusive, argument for ethical relativity. He examines the sorts of moral variations that do occur—for example, variations arising from differences of opinion and knowledge regarding the non-moral qualities of acts or their consequences; and variations arising from the possibility of alternative ways of satisfying primary needs. He also examines the nature of moral ideals and their relation to our basic needs. He appears to argue that we can proceed rationally to decide between differing moral judgments very much further than is commonly supposed—by submitting them to rational tests such as coherence and comprehensiveness. He says that " it may quite well be that advance in ethical thought has been hindered by the comparative failure of the social sciences to throw light on the relevant facts." But he concludes on a cautious note. "Whether any fundamental divergences in moral outlook will remain when the facts have been clarified and the ideals elucidated remains to be seen." However, one of the types of variation in morals is that "due to differences of moral insight and general level of development, moral and intellectual." Ginsberg attaches great importance to differences in the level of moral development, and this fits in with his argument about the nature of sociology. He maintains that the comparative method in sociology is a valid one, and that its further and extended use is of great importance. In using this method, sociology will have to face up to its immediate tasks of investigating the nature and criteria of development, and of arriving at an adequate social morphology. Some of the essays contain a number of suggestions about the analysis of the key concepts of social evolution and development.

I am inclined to think that Ginsberg is quite right in maintaining that we can get much nearer rational agreement over moral matters than we normally suppose. He does not deal, however, with the query : how is he rationally to persuade people to adopt an ethic that satisfies his tests of rationality ? If we are interested in achieving rational agreement over morals and in controlling the sources of disagreement, we will regard lack of knowledge about fact as an important and deplorable state of affairs. Unfortunately, we shall not be able to deal with this limitation to our capacities until much more social resources have been devoted to the investigation of fact than seems to be forthcoming at present and in the foreseeable future. Ginsberg's view of sociology is an unfashionable one in this period of specialisms and ultra-cautious empiricism But he is almost certainly right to insist that, if sociology is to be able to generalize at all, it will have to make use of the comparative method. He is also quite right to emphasise the developmental features of societies. For these are features that are just apt to be ignored by the current practitioners in the field, in spite of the large role they play in current popular and political judgments, and in spite of their intrinsic importance. What is especially memorable about these essays is what they reveal of Ginsberg's own moral ideals. They reveal, in particular, his belief in the value of sober and objective assessment of issues, and yet also his deep and passionate humanity, and his faith in the power of human reason to contribute to our social development. B. A. FARRELL.

KASTELL, JEAN (1962). *Casework in Child Care.* London : Routledge and Kegan Paul, Ltd., pp. xi+306, 35s.

In England and Wales there are more than sixty-thousand children in the care of local authorities. This book describes the statutory provision for their welfare, and the work of children's departments on their behalf. It contains little that is new, but it makes an excellent appraisal of the problems and methods of social casework and gives a clear picture of its difficulties and limitations. It carefully describes the way a child-care officer deals with such problems as taking a child into care, placing him with foster-parents and arranging for his adoption, and illustrates these procedures by detailed case studies.

The book is well written, in clear, unemotional language, and shows a deep concern for the welfare of deprived children. It shows a sympathetic understanding of their needs and those of the adults who work with them.

It is the best book on the topic known to the present reviewer. M. J. TYERMAN.

MAYS, J. B. (1962). *Education and the Urban Child.* Liverpool : Liverpool University Press, pp. ix+208, 25s.

The appearance of books and articles on the theme of the 'blackboard jungle,' has focussed attention on the problems of the school in the city slum. Dr. Mays and his colleagues in Liverpool have studied the question without sensationalism, have discussed a number of relevant specific issues and have related their observations to the wider theme of "the relationship between education and social change in the central area of a modern city" (page 1). In the concluding chapter, the aim is summarized as "to describe the work of a group of inner city schools in relation to the general social background and to assess their achievement in promoting the social objectives explicit and implicit in the 1944 Education Act " (page 178).

The factual material of the report provides useful information in a neglected area of educational research. It shows, for example, that the schools in this area are not a ' blackboard jungle,' but a reasonably organised educational provision with a fairly stable teaching force. As the evidence accumulates, we see more clearly what is involved in the gigantic task of establishing equal opportunity in areas of social disadvantage.

Specific issues are discussed from time to time : parent-teacher relationships, the gulf between the school curriculum and city life, the teacher as a welfare worker, the visiting of homes, the growth of the county college. Use of the index will locate a number of brief practical essays on important topics set among the factual material.

The general sociological reflections, consideration of the obstacles to equality of opportunity and the nature of the sub-culture of the area, are dealt with mainly in the first chapter and the last (which, perhaps, should be read first), but also in other chapters.

These three levels of writing on the subject are mixed throughout the text. According to their interests, readers will find some parts tedious, but they will also find pages to reflect on. Social factors in education have not even yet received the attention they deserve. Books on general sociological themes, or on anthropology, or social psychology, or sociometry, often lack a clear educational relevance. The kind of survey which Dr. Mays has carried out in Liverpool and the kind of report he has written make a useful contribution to the literature which teachers and those who train teachers should be familiar with. J. D. NISBET.

MURPHY, L. B., *et al.* (1962). *The Widening World of Childhood.* Basic Books, Inc., pp. xi+399, $10.00.

This is one of a number of publications arising from longitudinal studies of normal children. The book deals specifically with the presentation and initial analysis of straightforward observations on thirty-two pre-school children placed in a number of new situations over a period of time. Such situations included both structured and semi-spontaneous types. The authors have sought to discover something of the ways in which children of this age cope with strange situations until either they attain some degree of mastery or they relinquish both interest and attention. Thus, the work is concerned with behaviour analysis, but there is incidental reference to the ends attained by individual children.

Working from the observational material, the authors demonstrate the complex interplay of many factors determining the child's reactions to his environment, amongst which home influence, past experience, present perceptual and affective organisation and temperamental qualities are important. Newer from the point of view of psychological analysis are the ensuing attempts to show sequences of behaviour accompanying individual progress towards mastery. Parts III and IV of the book outline a comprehensive framework of such sequences, showing both variety and similarity in children of widely differing talents and backgrounds. In one sense, this framework represents a more detailed analysis of Piaget's concepts of *assimilation* and *accommodation*.

Studies of observational data fulfil an important research role, both in placing a perspective on information known to workers through more subjective and less organised means, and in supplying hypotheses for further study in a manner which is not possible in a more rigidly defined experimental situation. Though such studies are both time-consuming and difficult to carry out, each adds considerably to the body of knowledge about children, quickly repaying the effort involved. The present study offers a further store of material prolific in its interpretations for those concerned with children at any stage. Yet two advances have still to be made. In this investigation, it is pleasing to see the observation of ' normal ' children—for so often it is their ' abnormal ' counterparts. But the situations in which the children were observed are still very far from ' normal ' in any sense. More particularly, all analysis in the present work, as in all others known to the writer, is qualitative. Methods of modern statistics have not yet been adapted to furnish tools for an adequate appraisal and enrichment of such work. D. M. LEE.

PEAR, T. H. (1962). *The Moulding of Modern Man.* George Allen and Unwin, pp. 220, 21s.

If Professor Pear had not existed he must surely have been invented ! Someone of wide sympathies and great enthusiasms had to take care that all the subjects which intelligent laymen have been concerned about over the past half century should not be ignored just because intelligent psychologists have to be so nervously cautious in view of the equivocal scientific status of their discipline. But Professor Pear's interests, sensitivities and activities are so multifarious that it would have been a project for group research to assemble them in some sort of unity—and one can guess what would have happened if a committee had been responsible for the enterprise.

There is no question but that man is moulded by society and moulded by conditioning processes both in his formative childhood and in more active maturity, processes of which he is only in our century becoming scientifically aware. But early philosophers, sociologists and literary critics said as much long ago and supported the thesis by cogent argument and evidence. What makes us today so sensitive to Vance Packard and Madison Avenue ? Why do we fear the emergence of new magicians in the persons of brain-washers, sub-perception operators and sleep-indoctrinators, happiness and truth-drug pedlars ?

Might the psychologists have some responsibility ? They have certainly been insufficiently insistent upon one brute empirical fact, unexplained throughout the story of human enquiry and probably not susceptible to the process of explication, namely that human beings are in some crucial sense active and creative centres of that great energy of evolutionary process which ensures that 'novelty always arrives to enrich.' It is an emphasis on what *we* can do to control the forces that do so much to us and our children that makes Professor Pear's book so precious. But is discursive style and organisation and its offering of not very full or very well supported analyses of problems like indoctrination and opinion-making are sometimes impatient making. These problems are inexplicably bound up with value-judgments and the dramatic but very special modern cases in ' Grand Guignol ' dress have not changed the problems in principle. It was Freeman, in his Social Psychology, who exhibited as an example of the unconcsious 'moulding of man' Thorndike's Arithmetic of 1917. In the second book of this work in two hundred pages there are 643 problems which accept as a natural basis the profit motive and the structured private capitalism. Freeman suggested problems of the type of " If in a Southern cotton mill, one out of every hundred has pellagra, how many new cases can be expected if we increase the number employed by one thousand ? " But arithmetic texts containing such problems would be ' propagandist ' and not ' educational.'

The problems which Professor Pear reflects upon are so many and vary so much in size and disciplincary complexity that it is perhaps appropriate that the style is that of the essayist, humane and erudite but hurrying. We are hurried through the intellectual mazes of the Nature of the Self, the Relation of the Individual to Society, the Concepts of Loyalty, Treachery and the Definitions of Culture and the Reality of Progress, with occasional descents into factual reportage of the modern kaleidoscope of curious happenings and sensational writings about which it is so interesting and so easy for us all to read. Professor Pear makes it even easier, but one is troubled as to whether it is always quite a good thing. Is it not increasing the whole unhealthy bother about the ' hidden persuaders ' so that we forget our obligation to adopt values which require us to do some quite conscious thinking, some deliberate and avowed persuading and some strenuous and ingenious activity in terms of these values ? This energetic response from Professor Pear should, of course, stimulate something more active than a Grand Guignol ' frisson ' in his readers. GEORGE WESTBY.

SCHONELL, F. J., MCLEOD, J., COCHRANE, R. G. (eds.) (1962). ***The Slow Learner*: *Segregation or Integration*.** St. Lucia, Queensland, Australia: University of Queensland Press, pp. 92, 15s. 6d.

In this symposium—first published in 1960-61, in the University of Queensland's journal *The Slow Learner*—sixteen contributors from ten countries discuss four aspects of the education of subnormal children: the aims—social, scholastic and vocational; the relative efficiency, in achieving these aims, of various forms of special educational provision; the identification of those who would best profit by special education; the types of provision in different countries.

Within their 2,500 words, some chapters give a resumè of research; others are descriptive, statistical and historical, showing interesting discrepancies between policy in special education and general educational philosophy, for example, in England, U.S.A. and the U.S.S.R. The Russians' insistence on comprehensive diagnosis might well have come from a British textbook, but their reservation of 'remedial schools' for the organically subnormal raises many problems. Vernon, in a further contribution to his discussion of the I.Q., as a measure of potential, shows the difficulty of inference from psychometric data; in criticism of segregation he quotes both the evidence that some special school pupils excel their secondary modern counterparts, and the suggestion that segregation may increase retardation by lowering aspiration.

Several contributors claim that separate provision improves social integration by giving opportunity for guidance in social skills, and by eliminating "possibly the cruellest form of segregation—social rejection within the ordinary class." There is some supporting evidence that the subnormal child is happiest either in a special class or in a *small* normal class. A great limitation in discussion of this problem is the dependence even of experienced judgment on the *quality* of the provision with which the writer is familiar; well-controlled research is needed, on the lines of that carried out on gifted children. In the absence of such objective evidence, some of the descriptive contributions are naively insistent and repetitive, but Segal's five criteria for successful teaching in the ordinary school are soundly-based and well-supported.

The discussion is introduced and summed up by the deputy director of the remedial education centre at the Universy of Queensland The problem faced is not solved—nor is it claimed to be—but the reader will gain (as from any useful study in comparative education) protection from the risks of taking the familiar solution for granted, and of wasteful trial and error along lines already thoroughly explored elsewhere.

WILLIAM CURR.

SCHONELL, F. J., ROE, E., and MEDDLETON, I. G. (1963). ***Promise and Performance*.** University of London Press and University of Queensland Press, pp. xxv+405, 34s.

Francis Bacon wrote, "Some books are to be tasted, others to be swallowed, and some few to be chewed and digested." This book belongs without question to the last category, but after the mental digestion, it should be placed on one's reference shelf as a mine of valuable information. Though the book is concerned only with the University of Queensland and some of its findings have restricted applicability, there is much that has relevance to university problems in this country.

The enquiry studies the progress of 400 of the students over six years, examining their abilities, attainments and attitudes in relation both to their promise on entry and their subsequent performance. There are also sections on the functions of the staff, including lecturing and tutorials, with a brief but important comment on examining.

Mention is made here of a few points that are of particular interest. Even in this Australian university, with its relatively large intake, the working classes are severely under-represented, even when ability and attainment were held constant. A bi-serial correlation between matriculation score on the one hand and normal progress or failure to make normal progress, gave $r_{bis}=0{\cdot}48$. A valuable histogram

(p. 44) shows this relationship in detail. The impossibility of drawing any precise line which would permit almost all successful students to enter and at the same time exclude all potential failures is illustrated by the fact that in Matriculation Group IV one in every three students had made normal progress, while the proportion was one in four for Group V. In individual subjects a significant relationship between Matriculation results and first year university results was found in objective subjects such as Chemistry, Physics and Mathematics, but not in English, French, History or Latin. Only one-third of students who matriculated at the second attempt made normal progress later. (This and related findings reminds one that there will inevitably be a strong tendency for university markers to grade their students as good average, poor and failures *in accordance with the average standard of the student population in that course,* a standard which may change appreciably during a few years without altering the proportion of failures).

The relationship between I.Q. score and university performance was only 0·39. When I.Q. was added to matriculation score there was no improvement in the predictive power of the matriculation score, though students with both low matriculation and low I.Q. appeared to be a poor risk. Only 11 per cent. of this group graduated in minimum time, compared with 50 per cent. for other students. Whereas at Liverpool there was 71 per cent. normal progress, at Queensland it was 49 per cent. As one expects, there were wide variations in the failure rates of different Faculties, only 35·7 per cent of students making normal progress in Engineering. In all, only 22 per cent. of all students who failed in First Year (1950-52 entry), and 13·2 per cent. for the 1955 entry had no further failure.

One intriguing finding which does not surprise the reviewer but has many implications, is that although there was no significant difference between college and non-college students in their courses of study, the concensus of the evidence gives the non-college students the better performance—but readers need to make a detailed appraisal of the figures.

This is an important book which advances our knowledge. That it has not ' solved ' the problem of the selection of university students is not surprising, in view of the complexity of the problem. Indeed, as the authors realise, it can never be completely solved. The reviewer forecast some years ago that the ' break through ' will come in the field of temperamental charactistics and motivation, and the findings of the present book have strengthened this belief. R. R. DALE.

SHIELDS, J. (1962). *Monozygotic Twins Brought Up Apart and Brought Up Together.* Oxford University Press, pp. x+264, 50s.

This study follows the lines of the now classical inquiry by Newman, Freeman and Holzinger. But, instead of the nineteen pairs of twins in that inquiry, Dr. Shields has at his disposal forty-four pairs of monozygotic twins reared apart, a control group of forty-four pairs of monozygotic twins reared together, and a further group of thirty-two pairs of dizygotic twins, of whom eleven pairs were brought up apart. These twins were selected from volunteers after a B.B.C. appeal for twins in 1953. Though some bias in selection is recorded, for example, an excess of women and the middle aged, this does not invalidate the inquiry.

The investigation is reported carefully and objectively, and the conduct of it inspires confidence in the author's scientific integrity. There is no sound of axes being ground. It is all the more unfortunate that some aspects of the assessment are rather flimsy. The author is rather naive about intellectual assessment. " It might, therefore, be expected that a non-verbal intelligence test might be a good measure of innate capacity in twins, while any intellectual differences due to cultural or educational environment might show up more readily on the verbal intelligence test." (p. 58). It is a pity that more extensively used tests of intelligence were not used. The same point applies to the assessment of personality. Here Eysenck's Self-Rating Questionnaire for Neuroticism and Introversion–Extraversion has been used. While one of the most adequately standardised of these instruments, it also has a rather specific area of assessment. It is possible that the writer has recognised these difficulties, as he also develops a scheme of degrees of differences of personality based on personal interview. Perhaps, also, he overestimates the subjective element in such assessment. To this is added physical measurements and fully recorded case histories. These case histories occupy almost half the book, and make rich and interesting reading.

And the conclusions? The findings do not differ significantly from those of earlier studies in the same field. In the writer's own words, " 1.—Family environments can vary quite a lot without obscuring basic similarity in a pair of genetically identical twins. 2.—Even monozygotic brought up together can differ quite widely (p. 156). The interesting part is the way that there occur exceptions to the general trends, exceptions which have no apparent explanation. JAMES MAXWELL.

SPEARRITT, D. (1962). *Listening Comprehension—A Factorial Anaylsis.* A.C.E.R., pp. x+148, 42s.

The author of this book points out that as modern societies grow in size and complexity, the need for effective communication becomes all the more necessary, and although reading and writing are the more recently developed of the communication skills they seem to have been more readily accepted as subjects of school instruction than have the older skills of listening and speaking. Yet, we have hitherto had no factor analytic studies that have included listening comprehension tests. This book, which is (in effect) an extremely systematic and thorough Ph.D. thesis, not only makes an original contribution to our knowledge, about listening comprehension, but also serves as a model to research students, particularly in the factorial field, preparing to write an account of their own investigation. The techniques employed are in line with much recent American factor analytic practice, i.e., the use of normalised T scores, product moment correlations, principal components, oblique analytic rotation (Biquartimin) towards simple structure, and the extraction of factors with eigenvalues greater than unity, followed by second order factorisation. It is interesting to find that the human rotations give a poorer simple structure than the Biquartimin rotation, yielding hyperplanes of something like 10 per cent. less density, with very little compensating advantage in the number of relatively high loadings. The results, based on thirty-four tests, are most informative, as few of the hypotheses put forward by the author were fully verified. For example, the suggestion that there is no factor of listening comprehension separate from reasoning, verbal comprehension, attention and memory was rejected. Nor could the hypothesis be sustained that listening comprehension is related to meaningful memory but not to rote memory and span memory. Of course, the results do not imply that listening comprehension should be compulsorily taught simply because it has been shown to be an independent factor. The case for formal teaching of listening comprehension must be decided on its own merits.

There is little to fault in this book except perhaps the statistical appendix, which will surely be too elementary for the cognoscenti and too difficult for everyone else.

In summary, this book is recommended, not only for readers interested in listening comprehension, but also for those concerned with the design of theses on experimental topics, particularly in the area of factor analysis. F. W. WARBURTON.

PUBLICATIONS RECEIVED

J. B. Mays (1962). *Education and the Urban Child.* Liverpool University Press, pp. ix+208, 25s.

H. C. Lindgren (1962). *Educational Psychology in the Classroom.* (2nd edition). John Wiley and Sons, pp. xvii+475, 53s.

H. A. Witkin, *et al.* (1962). *Psychological Differentiation.* John Wiley and Sons, pp. xii+418, 60s.

Quinn McNemar (1962). *Psychological Statistics* (3rd edition). John Wiley and Sons, pp. vii+451, 59s.

Upton Sinclair (1962). *Mental Radio.* Springfield, Illinois : Charles C. Thomas, pp. xv+237, $8.50.

H. J. Eysenck (1962). *Know your own I.Q.* Pelican Book, pp. 192, 3s. 6d.

Mental Health Book Review Index, Vol. 7, 1962, Whole No. 12, pp. xiii+66. Sponsored by the World Federation for Mental Health and others. Annual subscription, $3.00.

Mary D. Ainsworth, *et al.* (1962). *Deprivation of Maternal Care.* Geneva : World Health Organisation, available through H.M.S.O., pp. 165, 12s.

G. R. Hawkes and D. Pease (1962). *Behaviour and Development from 5 to 12.* New York : Harper and Brothers, pp. xiv + 375, 46s.

Nirmala Kher, *et al.* (1962). *An Introduction to Child Development.* London : Asia Publishing House, pp. x + 174, 28s.

M. F. Cleugh (1962). *Educating Older People.* Tavistock Publications, pp. 185, 28s.

Harriet Wilson (1962). *Delinquency and Child Neglect.* George Allen and Unwin, pp. 195, 25s.

Glenn M. Blair, *et al.* (1962). *Educational Psychology* (2nd edition). New York and London : Macmillan Company, pp. xxiv+678, 53s.

Jean Kastell (1962). *Casework in Child Care.* Routledge and Kegan Paul, pp. xi +306, 35s.

H. R. Lantz and E. C. Snyder (1962). *Marriage.* John Wiley and Sons, pp. xii+ 417, 49s.

Southgate Group Reading Tests, Test 2 (Forms A and B). University of London Press, Ltd. (no price stated).

D. M. Armstrong (1962). *Bodily Sensations.* Routledge and Kegan Paul, Ltd., pp. vii+132, 12s. 6d.

M. V. C. Jeffreys (1962). *Personal Values in the Modern World.* Penguin Books, pp. 172, 3s. 6d.

Henry P. Smith (1962). *Psychology in Teaching* (2nd edition). Prentice-Hall, Inc., pp. xiv+498, 56s.

Albert J. Harris (1962). *How to Increase Reading Ability* (4th edition). Longmans, pp. xxi+624, 50s.

George G. Thompson (1962). *Child Psychology* (2nd edition). Houghton Mifflin Co., pp. xxi+714, 55s.

Fillmore H. Sanford (1962). *Psychology.* Prentice-Hall Inc., pp. 569, 42s.

W. A. Scott and M. Wertheimer (1962). *An Introduction to Psychological Research.* John Wiley and Sons, Ltd., pp. xiii+445, 60s.

Macpherson Lawrie (1962). *The Anatomy and Excellence of Human Nature.* The C. W. Daniel Co., Ltd., pp. 95, 7s. 6d.

F. J. Schonell, J. McLeod and R. G. Cochrane (edits.) (1962). *The Slow Learner.* University of Queensland Press, pp. vii+92, 15s. 6d.

E. L. Edmonds (1962). *The School Inspector.* Routledge and Kegan Paul, Ltd., pp. x+202, 28s.

Elizabeth Duffy (1962). *Activation and Behaviour.* John Wiley and Sons, Ltd , pp. xvii+384, 60s.

Edgar Anstey (1962). *Committees : How they Work and How to Work Them.* George Allen and Unwin, Ltd., pp. 116, 15s.

A. M. KEAN (edit.) (1962). *Researches and Studies.* The University of Leeds Institute of Education, pp. 103, 5s.

SIGMUND FREUD (1962). *Two Short Accounts of Psycho-Analysis.* Penguin Books, pp. 175, 3s. 6d.

C. A. MACE (1962). *The Psychology of Study.* Penguin Books, pp. 120, 3s. 6d.

MICHAEL HUTCHINSON AND CHRISTOPHER YOUNG (1962). *Educating the Intelligent.* Penguin Books, pp. 240, 3s. 6d.

LOIS BARCLAY MURPHY and Collaborators (1962). *The Widening World of Childhood.* New York : Basic Books, Inc., pp. xiv+399, $10.00.

JOHN MONEY (edit.) (1962). *Reading Disability—Progress and Research Needs in Dyslexia.* Johns Hopkins Press : London : Oxford University Press, pp. x+222, 40s.

VIDA CARVER (edit.) (1962). *C. A. Mace—A Symposium.* (Methuen and Co., pp. 134, 21s.

THELMA VENESS (1962). *School Leavers.* Methuen and Co., pp. xxv+252, 25s.

E. PHILIP TRAPP and PHILIP HIMELSTEIN (edits.) (1962). *Readings on the Exceptional Child.* Methuen and Co., Ltd., pp. xii+674, 55s.

JAMES J. JENKINS and DONALD G. PATERSON (edits.) (1962). *Studies in Individual Differences.* Methuen and Co., Ltd., pp. xiv+774, 63s.

JAMES SHIELDS (1962). *Monozygotic Twins.* Oxford University Press, pp. x+264, 50s.

DEREK RICHTER, *et al.* (edit.) (1962). *Aspects of Psychiatric Research.* Oxford University Press, pp. viii+445, 63s.

FRANK S. FREEMAN (1962). *Theory and Practice of Psychological Testing* (3rd edition) New York : Holt, Rinehart and Winston, pp. xviii+697, $7.95.

ARTHUR W. FOSHAY, *et al.* (1962). *Educational Achievements of Thirteen-year-olds in Twelve Countries.* Hamburg : Unesco Institute for Education, pp. 68, no price stated.

FREDERICK J. MCDONALD (1962). *Educational Psychology.* Prentice-Hall International Inc., pp. xix+748, 42s.

JOSEPH S. RUDLOFF (1962). *Recruiting Teachers for the Deaf.* Springfield, Illinois : Charles C. Thomas, pp. xii+68, $4.25.

REG. BUTLER (1962). *Creative Development.* Routledge and Kegan Paul, pp. v+82, 10s.,

JOHN APLEY and RONALD MACKEITH (1962). *The Child and his Symptoms.* Oxford : Blackwell Scientific Publications, pp. 262, 25s.

M. P. CARTER (1962). *Home, School and Work.* Pergamon Press, pp. xi+340, 50s.

MARY ATKINSON (1962). *Junior School Community* (2nd edition). Longmans, Green and Co., pp. xv+199, 13s. 6d.

JOHN C. DANIELS (1962). *Figure Reasoning Test.* Crosby, Lockwood and Son, Ltd., 4s. 6d. Instruction Manual, 1s. 6d., free with every 20 copies of test.

ROBERT T. OLIVER and DOMINICK A. BARBARA (1962). *The Healthy Mind in Communion and Communication.* Springfield, Illinois : Charles C. Thomas, pp. xi+166, $6.50.

DONALD SPEARRITT (1962). *Listening Comprehension—A Factorial Analysis.* Australian Council for Educational Research, pp. x+149, 42s. Aust.

F. Y. THOMPSON and A. C. SPENCE (1962). *English Second Language Course Bk.* 3, *Advanced.* Methuen and Co., Ltd., pp. 163, 10s.

TSUNG-YI-LIN and C. C. STANDLEY (1962). *The Scope of Epidemiology in Psychiatry.* Geneva : World Health Organization, pp. 76, 5s.

Mental Health Book Review Index, Vol. 7, 1962. Whole No. 12. With an editorial : The conservation of knowledge about human behaviour. xiii, pp. 66 Sponsored by the World Federation for Mental Health and others. Annual subscription : $3.00. The Index, c/o Miss Lois Afflerbach, Paul Klapper Library, Queen's College, Flushing, 67, N.Y., U S.A.

HIGHER EDUCATION IN THE UNITED KINGDOM—A Handbook for Students from Overseas and their Advisers. September, 1962, issued every two years. Longmans, Green and Co., pp. 244, 7s. 6d.

Hertha Riese (1963). *Heal the Hurt Child.* University of Chicago Press : London, pp. xxiii+615, 72s.

Fred J. Schonell, *et al.* (1963). *Promise and Performance.* University of London Press, pp. xxv+405, 34s.

Seonaid M. Robertson (1963). *Rosegarden and Labyrinth.* Routledge and Kegan Paul, pp. xxiii+216, 32s.

A. Stuart (1962). *Basic Ideas of Scientific Sampling.* Charles Griffin and Co., pp. 99, 18s.

S. M. Corey and J. K. Shukla (1962). *Practical Classroom Research by Teachers.* New Delhi : National Council for Educational Research and Training, pp. iv+116, 5s. 2d.

Elizabeth J. Levinson (1962). *Retarded Children in Maine.* University of Maine Press, pp. xii+265, no price given.

Clifford Allen (1963). *Passing Examinations.* Macmillan, pp. xiii+146, 12s. 6d

P. H. Mussen, J. J. Conger and J. Kagen (2nd edition, 1963). *Child Development and Personality.* New York and London : Harper and Row, pp. xiv+625, 56s.

Alton M. Broten (1963). *Houseparents in Children's Institutions.* Oxford University Press, pp. vi+90, 10s.

M. Metcalfe Smith, *et al.* (1962). *Research and Experiment in Education.* (Department of Education, King's College, Newcastle-upon-Tyne, pp. 67, 5s.

W. W. Cooley and P. R. Johnes (1963). *Multivariate Procedures for the Behavioral Sciences.* John Wiley and Sons, pp. x+211, 51s.

Sheldon Glueck (1963). *Law and Psychiatry.* Tavistock Publications, pp. ix+181, 38s.

Lee J. Cronbach (1963). *Educational Psychology* (2nd edition). Harcourt, Brace and World, Inc., pp. xxvii+706, $7.50.

Brian and Joan Simon (edits.) (1963). *Educational Psychology in the U.S.S.R.* Routledge and Kegan Paul, pp. xi+283, 40s.

B. Knapp (1963). *Skill in Sport.* Routledge and Kegan Paul, pp. xii+204, 30s.

H. M. Dresden-Coeners (1963). *The Psychology of Sex Instruction.* Sheed and Ward, pp. xv+139, 8s 6d.

George Humphrey (edit.) (1963). *Psychology Through Experiment.* Methuen and Co., pp. 307, 30s.

Stella E. Mason (edit.) (1963). *Signs, Signals and Symbols.* Methuen and Co., pp. xii+212, 35s.

Ronald Ridout (1963). *Better English*—An Introductory Book and Books 1 to 5, together providing a complete Primary course. Ginn and Co. : Intro.—3s. 9d., Bk. 1—3s. 9d., Bk. 2—4s., Bk. 3—4s. 6d., Bk. 4—6s., Bk. 5—7s.

CORRECTION

to article by Crookes and Green in the February issue.

Following line 7 on page 39 it should read: " . . . dominance. The difficulty in reading and writing would follow as a result of the . . . "

Some Important Books

THE DEVELOPMENT OF THE INFANT AND YOUNG CHILD
Second Edition. By Professor R. S. Illingworth, M.D., F.R.C.P., D.P.H., D.C.H.
370 pages, 126 illustrations. 30s.

THE NATURAL DEVELOPMENT OF THE CHILD
Fourth Edition. By Agatha H. Bowley, Ph.D., F.B.Psy.S.
224 pages, 84 illustrations. 10s. 6d.

THE PSYCHOLOGICAL CARE OF THE CHILD IN HOSPITAL
By Agatha H. Bowley.
48 pages, 8 illustrations. 4s. 6d.

CEREBRAL PALSY IN CHILDHOOD AND ADOLESCENCE
Edited by Professor J. L. Henderson, M.D., F.R.C.P.(Ed.). 411 pages. 35s.

E. & S. LIVINGSTONE, LTD.
TEVIOT PLACE, EDINBURGH

THE
BIRMINGHAM
PRINTERS LTD
PRINTERS & BOOKBINDERS

High Class
Technical
Publications

42-44 HILL STREET
BIRMINGHAM 5
TELEPHONE: MIDLAND 0196-7

OCCUPATIONAL PSYCHOLOGY

Editor: ALEC RODGER

OCTOBER, 1962 Volume 36, No. 4

CONTENTS

An Exploratory Study of Merit Rating Payment Systems in Three Factories. By R. MARRIOTT

Sources of Satisfaction in Repetitive Work. By ARTHUR N. TURNER and AMELIA L. MICLETTE.

Youth Employment Service Interviews: Part 2, Differences Between Interviews. By P. CAVANAGH, R. I. DRAKE and K. F. TAYLOR.

Other Books Received.

Book Review Index.

Annual Subscription 40 shillings.

National Institute of Industrial Psychology
14, Welbeck Street, London, W.1.

THE BRITISH JOURNAL OF MEDICAL PSYCHOLOGY

Edited by Dr. THOMAS FREEMAN

(*Contents of Vol. XXXVI, Part* 2, *June*, 1963)

FAIRBAIRN ISSUE

EDITORIAL.
W. R. D. FAIRBAIRN. Autobiographical Note.
J. D. SUTHERLAND. Object Relations Theory and the Conceptual Model of Psycho-Analysis.
OTTO F. KERNBERG. Discussion of J. D. Sutherland's paper.
H. V. DICKS. Object Relations Theory and Marital Studies.
J. C. B. SYM. Derivatives of Anal Erotism.
J. O. WISDOM. Fairbairn's Contribution on Object-Relationship, Splitting, and Ego-Structure.
H. GUNTRIP. Psychodynamic Theory and the Problem of Psychotherapy.
R. E. D. MARKILLIE. Observations on Early Ego Development.
J. F. PAVE. A Multi-Ego Theory of Libido-Destrudo Dynamics and Mental Illness.
REVIEWS.

Subscription price 75s. net per volume of 4 parts. Single parts 25s. net.

Orders may be placed with any bookseller, or sent direct to :

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200, Euston Road, London, N.W.1

PSYCHOLOGY AND RELIGION

G. Stephens Spinks

The first part of this book is devoted to the modern theories of the psychology of religion, particularly those of Freud and Jung, and the second half shows how such theories help to elucidate religious practices and the phenomena of mysticism. The author is a former editor of the Hibbert Journal. **28s.**

THE ART OF TEACHING

Gilbert Highet

This famous book is now available in the University Paperbacks series. **9s. 6d.**

METHUEN *& Co. Ltd. · 36 Essex Street · Strand · London W.C.2*

TAVISTOCK PUBLICATIONS

Pickford Projective Pictures

R. W. PICKFORD

With the assistance of Ruth Bowyer and John Struthers
Pictures by Winifred McElroy

This work, consisting of a set of picture material and an explanatory text, presents a new diagnostic and therapeutic instrument designed for use in child guidance work, in counselling, and in remedial teaching.

The picture material comprises 120 simple line drawings depicting children in various relationships with other children and with adults, and also sometimes alone. The accompanying text discusses the theoretical considerations underlying the design of the picture material and describes techniques for its administration.

Text 30*s. net. Picture material* 25*s. The set* 50*s.*

A Study of Children's Thinking

MARGARET DONALDSON

In collaboration with Donald Withrington
With an appendix by John Duthie

A study, based on systematic research, of the types of error commonly made by children in attempting to solve problems. The children studied fall into the age group 9-15 years, and the problems involved are of the kind commonly used in verbal group intelligence tests. In examining 'error' the author is concerned not merely with the 'wrong solution'; she has sought to discover the nature of the difficulties encountered by the children on the way toward solution.

The book makes a significant contribution to the theory of test construction as well as to the general topic of the development of intelligence. It is of fundamental importance for the teacher and for the educational psychologist. 35*s. net.*

A Textbook of Health Education

DENIS PIRRIE, M.B., B.Chir., D.P.H.
A. J. DALZELL-WARD, M.R.C.S., L.R.C.P., D.P.H.

With contributions by Winifred Warden, M.Sc., F.L.S., and D. Lynton Porter

The first comprehensive guide to the general principles of health education designed to provide teachers, students, and public health workers with a basic understanding both of methods and media of instruction and of the biological and social factors underlying problems of community health.

'The authors are to be congratulated on producing a text-book on health education practically free from jargon, and one which makes every topic they discuss quite clear.' *Royal Society of Health Journal.* 35*s. net.*

Educating Older People

M. F. CLEUGH

A study of the social and psychological factors that must be taken into account in teaching adults. Examples illustrating aspects of the teaching situation are discussed as a basis for the formulation of general principles in the education of the older student.

'Its findings are illustrated with detailed recordings of actual conferences and lecture groups and the whole is presented in an attractive and eminently readable fashion.' *Higher Education Journal.*

'Almost every page is rich with first hand observations and examples; and the whole volume is written in a lively and delightful style. Every university and college teacher should study it.' E. L. RIDLEY, *British Journal of Statistical Psychology.* 28*s. net.*

WILEY
Interscience

Concept Learning:
An Information Processing Problem

Earl B. Hunt, *Western Management Science Institute, University of California, Los Angeles.*

A particular topic, concept learning, has been examined from a problem oriented point of view. The research of the psychologist, the logician, and the designer of computing machinery is discussed as it relates to the topic. The problem chosen is certainly an important one. The ability to see objects as members of an abstract class of things is one of the most important talents people have. This monograph examines the problem of how any intelligence, human or artificial, can learn to assign objects to pre-existing classes. 296 *pages*, 57*s*.

Birth to Maturity:
A Study in Psychological Development

Jerome Kagan, *Fels Research Institute* and Howard Moss, *National Institute of Mental Health.*

Summarizes an intensive study of 71 adults who were members of the Fels Research Institute's longitudinal population. Each of these adults was intensively observed, tested, and studied from the time he was born until he was 14 years of age. Thus, there was available for each individual a large segment of rich and accurate information on his personality and intellectual development from birth to adolescence. 400 *pages*, 65*s*.

Computer Simulation of Personality

edited by S. S. Tomkins, *Princeton University* and S. Messick, *Educational Testing Service.*

Consists of the papers and the formal discussions of these papers presented at a conference on computer simulation and personality theory, held at Educational Testing Service and Princeton University in June, 1962. The purpose of the conference was to accelerate the dissemination of information concerning computer simulation among investigators in the field of personality and to achieve some evaluation of the potential of computer simulation for the field of personality. 325 *pages*, 40*s*.

Psychological Differentiation
Studies of Development

Herman A. Witkin, Ruth B. Dyk, Hanna F. Faterson, Donald R. Goodenough and Steven A. Karp, *all of the Department of Psychiatry, State University of New York College of Medicine.*

This book is concerned with deep-seated psychological patterns, encompassing the cognitive, personality and social functioning of the individual. 448 *pages*, 60*s*.

John Wiley · Interscience
Glen House · Stag Place · London SW1

OTHER PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology.
The British Journal of Psychology (Monograph Supplements).
The British Journal of Medical Psychology.
The British Journal of Statistical Psychology
The British Journal of Social and Clinical Psychology.

The British Journal of Psychology and *The British Journal of Medical Psychology* appear in quarterly parts in volumes of about 350 pages, subscription price of each, payable in advance, 60s. per volume (post free). Subscriptions may be sent to any Bookseller or to the Cambridge University Press, Bentley House, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $10.00. Enquiries should be addressed to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22.

The British Journal of Statistical Psychology is issued twice a year, in May and November, in volumes of about 184 pages. The subscription rate payable in advance is 30s. 6d. per volume (post free). The subscription price in the U.S.A. is $4.50 (post free). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, E.C.4., or to any Bookseller.

The British Journal of Social and Clinical Psychology is issued three times a year in February, June and November, in volumes of approximately 240 pages. The subscription price of each, payable in advance, is 50s. (post free). Subscriptions may be sent to any Bookseller, or to the Cambridge University Press, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $8.50. Enquiries should be sent to the Cambridge University Press American Branch, 32, East 57th Street, New York 22. Separate parts will be sold at 20s. : $3.50 in the United States.

A series of *Monograph Supplements* is issued in connection with the *British Journal of Psychology*, but is not included in the subscription for that journal. Orders may be sent to any Bookseller or to the Cambridge University Press, 200, Euston Road, London, N.W.1., from whom details of these publications may be obtained.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to :
Dr. BORIS SEMEONOFF, Department of Psychology, The University, Edinburgh.

For the *British Journal of Medical Psychology*, should be sent to Dr. THOMAS FREEMAN, Lansdowne Clinic, 4, Royal Crescent, Kelvingrove Glasgow, C 3.

For the *British Journal of Educational Psychology* should be sent to Mr. L. B. BIRCH, Institute of Education, The University, Sheffield, 10.

For the *British Journal of Statistical Psychology*, should be sent to the Assistant Editor, Sir CYRIL BURT, 9, Elsworthy Road, Hampstead, London, N.W.3.

For the *British Journal of Social and Clinical Psychology* should be sent to Mr. MICHAEL ARGYLE (Social Psychology Editor), Institute of Experimental Psychology, 1, South Parks Road, Oxford ;

or to :

Dr. JACK TIZARD (Clinical Psychology Editor), Maudsley Hospital, Denmark Hill London S.E.5.

For the *Monograph Supplements*, to Professor L. S. HEARNSHAW, Department of Psychology, The University, 7, Abercromby Square, Liverpool, 7, whose decision regarding the publication of all *Monograph Supplements* will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1st September of the preceding year.

OFF-PRINTS

Contributors receive 50 free copies of their article and may buy more in units of 50.

Members of the British Psychological Society receive all the Society's Journals at reduced rates.

Information on membership and further activities may be obtained from the Secretary, Tavistock House South, Tavistock Square, London, W.C.1.

PUBLISHED BY THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May and distributed free to members of the Association. To non-members it costs, 5s. per issue, post free (subscription 14s. per annum).

Members of the A.T.C.D.E. receive the *British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151, Gower Street, London, W.C.1.

CONTENTS

		PAGE
R. S. MacArthur and W. B. Elley :	The Reduction of Socioeconomic Bias in Intelligence Testing	107
L. Hudson :	The Relation of Psychological Test Scores to Academic Bias	120
C. S. Rushton and A. E. Stockwin :	Changes in Terman-Merrill I.Qs. of Educationally Sub-Normal Boys	132
G. Jahoda :	The Development of Children's Ideas about Country and Nationality : II.—National Symbols and Themes	143
A. S. Phillips :	The Self Concepts of Trainee Teachers in two Sub-cultures	154
J. D. McNeil and E. R. Keislar :	Value of the Oral Response in Beginning Reading : An experimental Study using Programmed Instruction	162
J. D. Nisbet and R. Illesley :	The Influence of Early Puberty on Test Performance at the age of eleven	169
G. W. S. Mackay and P. E. Vernon :	The Measurement of Learning Ability	177
Critical Comments : T. H. B. Hollins J. M. Thyne	Teaching Educational Psychology in Training Colleges	187
Comment : H. J. Eysenck	On the Relation of Neuroticism and Extraversion to Intelligence and Educational Attainment . .	192
Critical Notices : P.E. Vernon	Intelligence and Experience by J. McVie Hunt . .	193
Cyril Burt	The Appreciation of Beauty by C. W. Valentine . .	194
Book Reviews		202
Publications Received		210

10 AUG 1963

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

CALCUTTA
UNIV

NOVEMBER 1963

VOL. XXXIII, PART 3

Price One Pound

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
L. B. BIRCH

Associate Editors:
P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL

SIR FRED J. SCHONELL
C. W. VALENTINE
S. WISEMAN

Vol XXXIII, Part 3

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education, once a term, in February, June and November. The subscription is £2 a year post free payable in advance, or £1 each number.

Subscriptions and Orders should be sent to "The Manager, B.J.E.P. Department, Methuen and Co., Ltd., 36, Essex Street, Strand, London, W.C.2." Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to L. B. Birch, The Institute of Education, The University, Sheffield, 10. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to L. B. Birch, at the above address. *Advertisements* and *Subscriptions* for this *Journal* should be sent to Manager, B.J.E.P. Department, Methuen and Co., 36, Essex Street, Strand, London, W.C.2.

Published by
METHUEN AND CO., LTD., 36, ESSEX STREET, STRAND, LONDON, W.C.2.,
for the British Journal of Educational Psychology, Ltd.

Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

VOLUME XXXIII NOVEMBER, 1963 PART 3

TEACHERS' PERSONALITY DESCRIPTION OF SIX-YEAR-OLDS: A CHECK ON STRUCTURE

By RAYMOND B. CATTELL
(*University of Illinois*)

SUMMARY. Since factor analytic evidence on personality structure in young children is still uncommon and incomplete, an independent, checking analysis is presented on the excellent data gathered by Digman, on 102 seven and eight year-olds, by use of behavior ratings on a standard set of thirty-nine bipolar ' marker ' traits.

The general agreement with Digman's findings is good, seven of his eleven factors closely matching seven of the twelve here, in ' blind ' rotations. However, where there are differences, the technical points by which differences in factor analyses may be evaluated are illustrated. In particular, the differences of interpretation are evaluated, for the present writer's factors A, cyclothymia; C, ego strength; D, excitability; E, dominance; and so on to O, guilt proneness, are more exactly confirmed here. This better confirmation coincides with (*a*) a higher hyperplane count (mean 70 per cent.), and (*b*) a higher statistical significance of simple structure by the Bargmann test. However, the differences of the independent analyses are small compared with the agreements.

The study has relevance also to the cross-cultural constancy of personality factors, in that the subjects were roughly half Japanese, half Caucasian, in Hawaii.

I.—THE PRESENT STATUS OF RATING FACTOR RESEARCH.

SINCE Burt's (1940) pioneer factor analyses of ratings of personality traits in children, the main concepts have been considerably consolidated. Studies agree on some twelve to eighteen source traits (Cattell, 1957) which have been indexed as A through O, and include such dimensions as cyclothymia-vs-schizothymia (A), ego strength-vs-general emotionality (C), dominance-vs-submissiveness (E), surgency-vs-desurgency (F), and super ego strength-vs-super ego weakness (G). Although the popularly expressed notion that personality will be ' simpler ' in children may be in some sense true, the inference that it implies markedly fewer factors has not proved true. For, as the collations of results over age by Cattell (1957) have shown, the main factors show continuity from pre-school child to adult, the former being classifiable on such dimensions as dominance-submissiveness, cyclothymia, surgency, etc., as definitely as adults.

On the other hand, agreement among the three media of observation—L-data (life record, rating *in situ*), Q-data (questionnaire) and T-data (objective, situational or laboratory test)—has not been so clear. The present writer's conclusion (1961) is that the same number and nature of factors appear in L-data as Q-data, but that correlations between results from the two sources of scores are much reduced by intrusion of ' instrument factors ' and ' refraction factors,' especially in the rating field, e.g., between parents' and teachers' ratings. However, most will agree with such writers as Digman (1963), Pascal and Jenkins (1960) and Sells (1961) that clarification of the rating factor situation is important, because these personality factors are the real, *criterion* behavior, from which Q and T-data, as merely test data, must draw their validity.

Unfortunately, there are still some disagreements on the nature of personality factors in rated everyday life behavior which spring from the failure of many studies to reach what have long been recognized as minimal satisfactory research conditions. As discussed elsewhere (Cattell, 1957) these include: (*a*) use of more than one rater; (*b*) a position of the rater to permit observation of substantial varieties of behavior; (*c*) clearly, behaviorally-defined traits; (*d*) carrying over of a sufficiency of factor marker traits from study to study; and (*e*) thorough factor analytic technique with respect to testing for number of factors and rotating to an unquestionably unique simple structure resolution.

Since the last survey made by the present writer (1957) only three of the accessible studies came close to meeting these requirements, namely, those by Tupes and Cristal (1958), Norman (1959) and Digman (1963). The former suffer from under-extraction, but they nicely confirm the factors concluded from earlier work, as far as they go. The systematic studies by Digman are of particular interest, and will be made the basis of analysis here, because (*a*) the data are very soundly gathered; (*b*) they contribute to an age level—the 6 to 8 year old level—on which few clear findings exist, except those of Koch (1942), Peterson and Cattell (1959), and Cattell and Coan (1957), and which, because of the formative importance of these early school years, are of especial importance to teachers; and (*c*) they are gathered in a culture—Hawaii, with about half Japanese families—sufficiently different from the results previously collated by Cattell (1957), to test the hypotheses hitherto checked only by Q-date (Cattell, Pichot and Rennes (1961), Gorsuch (1963), Tsujioka (in press) that the personality factors A through O are so basic in personality structure that they essentialy transcend cultural and racial differences.

II.—METHODOLOGICAL STANDARDS OF A RE-ANALYSIS.

The fact that Digman's research, in spite of its excellent foundations with respect to requirements (*a*) through (*d*) above, has only somewhat more than half its factors agreeing with those of the collated results from other sources, justifies a special demonstration of the importance of steps in (*e*). This article proposes an independent analysis of Digman's raw data; a sharp focussing of the technical points at issue in the reaching of consistent findings in this field, and a crystallizing of the conceptions which can be used and tested in personality research with young children.

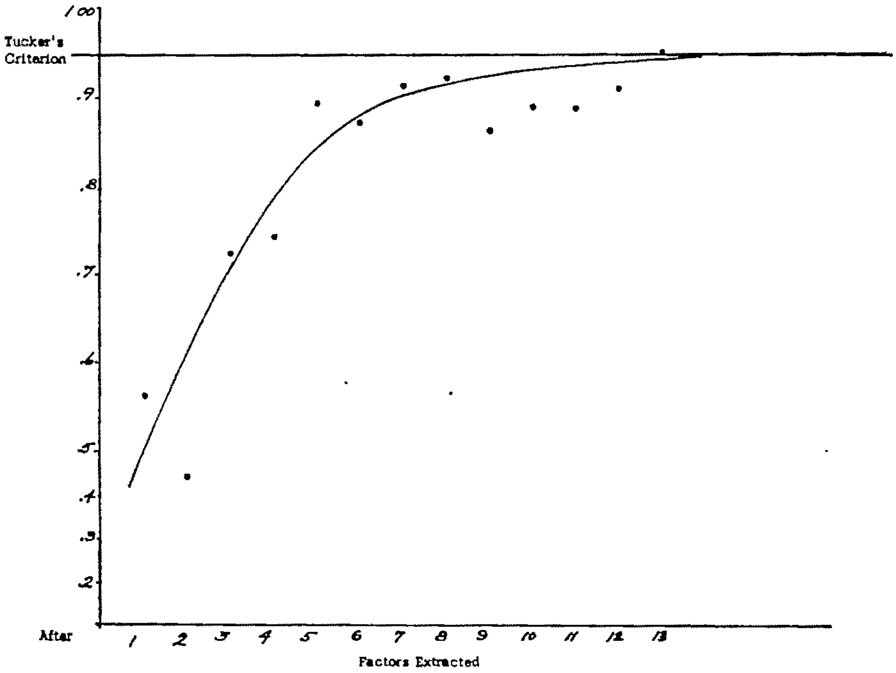
Digman began by using the comprehensive list of behavior variables distilled from many previous researches (Cattell, 1957) and which has been used by Coan and others as a marker framework. These thirty-nine bipolar defined variables were rated by teachers on 102 seven and eight year old boys and girls, by procedures which may be examined in Digman's account (1963). He used a principal axes solution, with estimated communalities, and experimented with an orthogonal varimax solution which, as Cattell and Dickman (1962) have tried to show, is incompatible with a truly maximal simple structure solution. The solution we present here differs from Digman's in (*a*) concluding that there are slightly more factors, and (*b*) rotating to a simple structure which ***maximizes the number of zeros***, regardless of whether the final solution presents uncorrelated or correlated factors. As a general principle, the present writer maintains that orthogonality occurs only as a rare and accidental feature when simple structure is single-mindedly pursued. Digman later moved to a final oblique resolution.

As tests of (*a*) above—completeness of extraction—we applied to the factorization of Digman's correlation matrix three tests, namely, those of

Tucker (1938), Sokal (1959) and the Kaiser-Pawlik method.* These converged on twelve factors. (See Diagram 1.)

DIAGRAM 1

EVIDENCE ON COMPLETENESS OF FACTOR EXTRACTION · TUCKER'S METHOD.



Communalities estimated by Burt's method (1940) were then inserted and, through five iterated factor analyses, were found to converge to fixed values for these twelve factors. The resulting unrotated matrix is shown in Table 1.

The search for simple structure was pursued in three steps of successive refinement: (1) by the Oblimax analytical computer program (Pinzka and Saunders, 1954); (2) by the correction of certain systematic faults in analytical programs through a tidying by the Maxplane program (Cattell and Muerle, 1960); (3) by a final hand polishing, maximizing the precision of the already clear hyperplanes by the Rotoplot (oscilloscope plot) program. Although the last is by hand and eye, it is kept, in the true scientific sense, 'blind,' since variables are unnumbered. The investigator thus cannot let his prejudices guide him, and the end result is a truly independent piece of evidence on personality factor structure.

---

* Tucker's method has hitherto been most used by the present writer, but it disregards the size of sample, and Sokal has shown (1959) that his method gives better results on empirical data for which the number of factors is known. Pawlik's modification of Kaiser's method, worked out in this laboratory, is to insert ones for communalities, to plot sizes of successive latent roots, and to stop (usually in the neighborhood of a latent root of unity) when a point of inflexion occurs as the descending curve becomes a straight line. It is the best method we have yet found, in agreeing with previous known answers.

TABLE 1

PRINCIPAL AXES FACTORS.

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1	46	65	−06	42	10	00	02	−10	−09	−05	12	−12
2	−71	09	−18	−01	19	03	−08	11	17	08	−07	−22
3	−42	10	−25	19	−27	−21	−18	−25	03	−05	04	21
4	34	83	−10	05	−03	04	−19	−01	−04	−16	06	−02
5	56	17	58	03	−25	−16	10	−12	−12	−01	−06	02
6	−68	27	36	05	15	−18	26	−00	19	−00	−06	07
7	70	39	−02	12	06	−33	06	08	12	17	−02	−01
8	69	−25	10	−26	−12	02	−05	11	03	26	06	01
9	−22	70	28	26	03	−11	−11	11	−02	−07	−08	−22
10	19	42	48	−37	−21	08	−04	10	13	−04	05	−05
11	42	−27	52	−15	−07	01	−01	−20	−04	−13	−11	−05
12	−54	42	−14	24	−08	−16	16	19	−10	−03	−12	10
13	73	−10	47	01	−09	23	03	04	10	07	02	−05
14	58	60	06	09	10	14	−18	04	10	−18	08	09
15	70	−40	30	15	09	04	29	−04	−15	−14	03	03
16	−72	−21	16	29	01	03	−00	09	−20	−01	01	−03
17	−71	17	35	17	07	11	00	−26	−01	07	03	09
18	−47	69	10	14	−04	13	−07	−08	21	14	−15	06
19	75	25	−38	−21	04	−08	−01	−10	03	−00	−18	04
20	−66	−11	26	28	05	26	00	04	−17	10	−05	08
21	−72	27	20	−09	−15	−11	09	16	02	12	15	07
22	−41	51	29	−25	36	−01	−14	−08	−10	−08	00	08
23	14	70	16	18	−09	−07	17	−02	11	−11	04	03
24	60	34	21	−07	28	07	01	09	21	02	−09	08
25	38	48	−28	−39	13	11	32	−11	−10	06	−15	00
26	72	−21	38	26	−03	08	−02	02	−02	20	01	−12
27	56	55	−35	−10	08	−12	12	−07	−14	−02	−03	−02
28	−21	80	13	−28	−16	11	−06	−06	−01	02	−05	04
29	−38	24	39	−33	36	−11	−13	09	−25	01	−00	02
30	−74	−03	19	10	05	−21	26	−09	19	−01	−03	−13
31	71	04	−02	17	02	04	−02	20	02	−14	−16	06
32	36	−23	19	29	−05	07	−03	23	04	−09	−21	17
33	−38	82	04	−08	−12	10	−15	10	−04	01	−05	03
34	04	38	−07	33	06	−06	−23	−31	−01	13	−27	−01
35	−18	65	−21	13	−32	21	10	10	−14	−01	−13	−11
36	−59	−65	−04	−19	−02	21	09	−00	02	−10	−14	04
37	61	23	08	30	14	−21	−04	−00	−13	19	104	10
38	−16	82	−13	−04	−01	17	20	−06	−11	21	03	−04
39	20	13	−28	42	19	36	12	−13	14	−01	15	07

TABLE 2

TRANSFORMATION MATRIX.

	1	2	3	4	5	6	7	8	9	10	11	12
1	— ·16	— ·12	— ·07	·07	·01	— ·10	— ·20	— ·08	— ·05	·01	— ·04	·20
2	— ·21	·08	— ·17	·03	·11	·03	— ·10	·33	·17	— ·03	·05	— ·03
3	— ·14	·24	·02	·10	— ·01	— ·15	·30	— ·05	·34	— ·05	·20	·45
4	— ·43	·08	·29	·18	·26	·03	·31	— ·14	— ·31	·01	— ·33	— ·20
5	— ·12	— ·03	·09	·11	·34	·16	— ·16	— ·51	— ·17	— ·17	·85	— ·04
6	·45	— ·28	·13	·16	·45	·05	·17	·40	— ·18	— ·08	— ·05	·42
7	·05	·57	·18	— ·13	— ·10	— ·42	— ·36	·40	— ·47	— ·10	·11	·02
8	— ·04	— ·35	·40	— ·42	— ·50	·05	·36	·27	— ·03	— ·44	·11	— ·16
9	·22	·51	— ·58	·28	— ·05	·52	— ·05	— ·15	— ·36	— ·04	— ·12	·31
10	— ·42	— ·12	·31	— ·34	·35	·42	·16	·37	— ·37	·33	·08	·35
11	— ·27	— ·32	— ·63	— ·49	·47	·44	·62	— ·12	— ·37	·06	— ·04	— ·05
12	·35	·02	— ·05	·53	·03	— ·34	·19	·16	— ·28	81	·27	— ·54

PERCENTAGE OF VARIABLES IN THE HYPERPLANES OF THE REFERENCE VECTORS.

Factor	No	Percent.
1	27	70
2	32	82
3	26	67
4	24	62
5	31	80
6	27	70
7	23	59
8	30	77
9	23	59
10	27	70
11	31	80
12	26	67
TOTAL	327	70

It is also methodologically important to point out that: (*a*) a simple structure resolution draws its objectivity not from machine handling, but from the fact of a unique maximum having been attained in the hyperplane count. In the present case a very thorough search through fourteen overall rotations after the oblimax and maxplane machine programs was carried out with results shown in Table 2. It will be seen that a mean value of 70 per cent. in the hyperplane was reached, with no single factor below 59 per cent.; (*b*) that being objective and maximal is not enough. It must also be significant. The significance of the simple structure was tested here by Bargmann's test (1954). This test, as argued elsewhere (Cattell and Pawlik, in press), needs improvement since it makes assumptions which in some situations result in under-estimation of significance. By this criterion one factor, No. 3, is non-significant and two, Nos. 5 and 8, are only at the 10 per cent. level, but the rest are mainly at $<\cdot01$ or $<\cdot001$. The rotation by Digman, which we are compelled to consider a less acceptable alternative interpretation of the factors to that given here, has, by the same criterion, four non-significant (His Nos. 1, 2, 3 and 4) two at ·25 (His Nos. 7 and 8) and the remaining five at the same level as our acceptable nine.

TABLE 3A

PRIMARY FACTOR PATTERN.

Variables	1	2	3	4	5	6	7	8	9	10	12	12
1	−67	−03	−06	−07	45	−11	−10	01	−05	−10	−10	−02
2	06	07	02	−07	−07	56	06	−01	−01	−28	10	−05
3	04	11	−12	−07	08	04	14	−04	01	39	−35	−60
4	−26	−19	−31	10	30	−03	−13	20	26	−07	−09	−08
5	−37	27	04	00	−13	−39	−06	12	34	09	−10	42
6	−03	73	−02	13	−08	04	08	09	01	−03	30	−06
7	−67	14	−14	06	−21	09	−32	−01	−11	02	03	06
8	−12	−30	−07	−07	−30	−00	−00	09	−07	09	−01	50
9	−51	16	01	10	05	16	14	12	40	−32	−02	06
10	07	10	−38	06	−13	−10	10	38	41	−12	02	59
11	12	19	02	03	01	−20	−16	−27	35	−00	03	56
12	−12	21	18	18	−18	09	06	41	00	−01	−06	−62
13	−13	−01	−04	15	03	−11	11	06	01	−07	−05	80
14	−15	−16	−34	37	38	−06	−02	−00	10	−03	05	09
15	−10	09	18	10	14	−56	−11	−26	−15	−03	09	24
16	−04	−03	36	−03	04	−07	56	−01	05	−08	−01	−26
17	−01	34	12	−09	33	06	38	07	06	18	11	06
18	−07	34	−04	19	09	43	08	45	07	09	−02	09
19	−00	−90	−10	07	−09	10	−82	01	00	09	01	−02
20	04	−01	43	10	14	−00	59	20	−02	05	08	−03
21	−06	19	−11	−13	−22	−07	48	43	06	02	03	−15
22	03	04	−11	06	25	01	02	−06	40	−01	56	−05
23	−29	39	−22	13	14	−17	−10	30	05	−03	−08	00
24	−13	11	−16	42	07	13	−27	−03	−02	−06	34	38
25	13	10	02	−13	−01	−09	−89	48	−04	−00	26	11
26	−50	−10	18	01	−00	−01	19	−12	−07	−05	−10	65
27	−28	−07	−09	−12	05	−14	−68	17	01	−00	06	−20
28	07	09	−21	−01	03	04	−10	61	39	05	02	15
29	−09	−13	06	−04	−03	−07	14	−10	47	−09	63	00
30	04	70	−01	−12	−05	07	02	−13	−01	−16	05	−08
31	−06	−13	06	51	−07	−09	−21	−06	−01	−10	−05	02
32	06	−01	22	64	−14	−07	16	−06	−04	01	−08	00
33	−04	−06	−12	12	01	13	14	58	35	−02	−01	−09
34	−18	04	39	−04	05	46	−16	−49	01	14	−17	08
35	−06	−04	15	00	−04	−00	−10	80	13	15	−35	−11
36	76	09	22	07	−10	03	03	−03	−05	−01	05	−05
37	−75	−11	11	04	03	−09	02	−13	−11	16	09	−05
38	−25	07	04	−25	14	03	−19	75	−05	02	05	05
39	−04	−03	−02	12	62	03	06	01	−58	07	−07	−03

INTERCORRELATIONS AMONG PRIMARY FACTORS.

	1	2	3	4	5	6	7	8	9	10	11	12
1	—	22	40	−62	−24	21	49	−22	−09	−16	12	−16
2		—	24	−38	18	33	63	20	34	−22	24	−40
3			—	−27	−01	03	39	−30	−10	−07	−07	−02
4				—	10	−38	−48	−04	−11	17	−30	43
5					—	34	−05	51	32	−12	05	−29
6						—	35	41	27	−19	26	−63
7							—	01	23	−23	28	−31
8								—	39	−20	31	−42
9									—	09	22	−25
10										—	−12	39
11											—	−22
12												—

TABLE 3B

LIST OF VARIABLES IN TABLE 3A

1. Non-aggressive, kind, considerate *versus* aggressive, tends towards fighting, bullying, teasing, cruelty.
2. Unpopular, generally disliked by other children *versus* popular, generally liked by other children.
3. Poor general health, prone to absence by reason of illness, or physical complaints *versus* of generally good health.
4. Quiet *versus* noisy, distracting in class
5. Free of jealousy, feels accepted *versus* prone to jealousy.
6. Learns slowly *versus* learns fast.
7. Conscientious, trustworthy *versus* untrustworthy, dishonest.
8. Self-sufficient, independent *versus* dependent on teacher.
9. Submissive, follows lead of other children *versus* self-assertive, tends to dominate other children.
10. Prefers not to be noticed *versus* demanding of teacher's attention.
11. Well poised, tough, sticks up for own rights when threatened *versus* easily upset, overwhelmed by teasing of other children, yields easily to persuasion
12. Lacking in self-confidence, easily discouraged or defeated *versus* confident (perhaps over-confident) of own ability and ideas.
13. Placid, free from distress *versus* fearful, worrying, anxious.
14. Calm, relaxed *versus* over-active, excitable, perhaps irritable.
15. Cheerful *versus* depressed.
16. Changeable in interests, attitudes, opinions *versus* stable in interests, attitudes, opinions.
17. Inattentive, absent-minded, prone to day dreaming, shows poor concentration *versus* very attentive to class proceedings.
18. Lethargic, apathetic, easily tired or fatigued *versus* alert, wide-awake, energetic.
19. Responsible *versus* irresponsible, frivolous.
20. Quitting, fickle *versus* persevering, determined.
21. Socially awkward and clumsy *versus* polished in manner
22. Practical-minded *versus* imaginative.
23. Self-abasive, deferent, minimizes own importance *versus* self-centred, conceited, boastful, ' show-off '
24. Lacks nervous habits *versus* prone to nervous habits (e.g , thumb-sucking, nail-biting, pulling and twisting hair, grimacing).
25. Neat, tidy, orderly *versus* untidy, careless with respect to appearance, of self belongings.
26. Adaptable, flexible *versus* rigid, has difficulty adjusting to changes or new situations.
27. Careful with property of others *versus* careless, destructive of property of others.
28. Secretive, reserved *versus* expressive, frank.
29. Lacking in artistic feeling *versus* aesthetically sensitive, aesthetically fastidious.
30. Has difficulty following instructions *versus* follows instructions easily and accurately.
31. Parents appear generally accepting toward child *versus* parents appear generally rejecting toward child.
32. Parents appear relatively permissive, allow child to make independent decisions *versus* parents appear dominating, over-protective toward child.
33. Shy, bashful, seclusive, aloof, remains fairly isolated from other children *versus* outgoing, mixes freely with other children.
34. Associates mostly with children of opposite sex *versus* associates mostly with children of own sex.
35. Prefers solitary pursuits *versus* gregarious, prefers games involving many children.
36. Negativistic, stubborn, disobedient, argumentative *versus* co-operative, compliant, obedient
37. Trustful of others, readily accepts solicitude of others as sincere *versus* suspicious of others, ungrateful, rejects affection or solicitude
38. Retiring, cautious *versus* adventurous, bold, willing to take the chance of possible rejection or injury.
39. Female *versus* male.

Possibly the hyperplane on our unsatisfactory factor No. 3 will be improved by some investigator trying further shifts, but perhaps also it is the nature of Ego Strength (see Table 6) that it affects too much of the personality to leave any marked hyperplane. At any rate its match to the pattern obtained with the maximum simple structure position in other studies is good. The accepted unique resolution for all factors is shown in Table 3A, but, for convenience, the labels for the variables (positive pole to the right) in it are listed in Table 3B. It should be noted that Table 3A is not left in the reference vector system (in which the ±·10 hyperplane count was taken) but in the *true* factor system, and the intercorrelations given below are similarly those of the primary factors, and are greater than for the reference vectors.

III.—INTERPRETATION OF FACTORS.

It is our purpose now to compare the meanings of the factors here obtained with those of previous studies. For convenience their order has been made the same in Table 3 as that of the letter index system used in dozens of other published studies in this area, and the axes have been reflected to put the positive signs in the familiar direction. Only the top six to eight variables are listed, corresponding to those clearly outside the hyperplane, and for the reader's convenience, the verbal labels have been adjusted to signs, i.e., by reading straight down the right-hand side one will get the labelled behaviors as they go together.

TABLE 4

FACTOR 2—SIZOTHYMIA *versus* AFFECTOTHYMIA, A.

Item No.	Factor Pattern	Trait Description
36	·76	Negativistic, stubborn, disobedient, argumentative *versus* co-operative, compliant, obedient.
37	—·75	Suspicious of others, ungrateful, rejects affection or solicitude *versus* trustful of others, readily accepts solicitude of others as sincere
1	—·67	Agressive, tends toward fighting, bullying, teasing, cruelty *versus* non-aggressive, kind, considerate.
7	—·67	Untrustworthy, dishonest *versus* conscientious, trustworthy.
9	—·51	Self-assertive, tends to dominate other children *versus* submissive, follows lead of other children
26	—·50	Rigid, has difficulty adjusting to changes or new situations *versus* adaptable, flexible.
5	—·37	Prone to jealousy *versus* free of jealousy, feels accepted.
23	—·29	Self-centred, conceited, boastful, 'show-off' *versus* self-abasive, deferent, minimizes own importance.

The markers for cyclothymia-vs.-schizothymia, namely co-operative, trustful and adaptable, occur here and nowhere else. This, together with its large size (A tends to rank as the first factor in size, as its original symbolization indicates) defines it. The only problem here is whether the Protension factor L, is mixed with this (L is otherwise missing) and might have separated if an extra dimension had been extracted. The s index of similarity puts the match of this with A in the Coan (1957) study beyond P < ·01 value. We believe sizothymia-vs-affectothymia would, in the future, be a better label for this factor.*

TABLE 5

FACTOR 2—INTELLIGENCE, B.

Item No.	Factor Pattern	Trait Description.
6	·73	Learns slowly *versus* learns fast
30	·70	Has difficulty following instructions *versus* follows instructions easily and accurately.
23	·39	Self-abasive, deferent, minimizes own importance *versus* self-centred, conceited, boastful, ' show-off.'
17	·34	Inattentive, absent-minded, prone to day dreaming, shows poor concentration *versus* very attentive to class proceedings.
18	·34	Lethargic, apathetic, easily tired or fatigued *versus* alert, wide-awake, energetic.
8	— ·30	Dependent on teacher *versus* self-sufficient, independen.
5	·27	Free of jealousy, feels accepted *vreus* prone to jealosy.
12	·21	Lacking in self-confidence, easily discouraged or defeated *versus* confident.

This is clearly Spearman's ' g,' picked up in ratings, as it often is, with the associated traits we are accustomed to find with it in the classroom situation, e.g., lack of self-confidence and inattentiveness at the low intelligence pole. The symbol B has been used for ' g ' in the total personality expression, reserving the latter for the strictly cognitive measurement.

---

* Because of its excellent confirmation of the Bleuler and Kretschmer schizothymia-cyclothymia dimension, Factor A has been so named since its discovery in 1945 (Cattell). Unfortunately, teachers and parents have associated the term schizothymia more with the schizophrenic abnormality which can attack schizothymes than with the normal and effective temperament. After fifteen years of battling with this misapprehension, we are suggesting that the A factor axis be relabled as *sizothymia-vs-affectotothymia* (or cyclothymia). This stresses the emotional detachment or flatness of A(–) (sizo from *assidere*, as in the root for painter's size, to make colors ' lie flat ') At the opposite pole, since in the normal range the *affective* aspect of this temperament is more prominent than the cycloid mood swings, it would improve the A + reference to call it *affecto-thymia* (agreeing, in abormal projections, with the currently favored designation ' affective ' rather than ' cyclic ' for the affective psychoses). The clearer sound distinction of sizothymic and affectothymic would also be a contribution to oral discussion,

TABLE 6

FACTOR 3—GENERAL EMOTIONALITY *versus* EGO STRENGTH, C.

Item No.	Factor Pattern	Trait Description.
20	·43	Quitting, fickle *versus* persevering, determined.
34	·39	Associates mostly with children of opposite sex *versus* associates mostly with children of own sex.
10	—·38	Demanding of teacher's attention *versus* prefers not to be noticed.
16	·36	Changeable in interests, attitudes, opinions *versus* stable in interests, attitudes, opinions.
14	—·34	Overactive, excitable, perhaps irritable *versus* calm, relaxed.
4	—·31	Noisy, distracting in class *versus* quiet.
23	—·22	Self-centred, conceited, boastful, 'show-off' *versus* self-abasive, deferent, minimizes own importance.
32	·22	Parents appear relatively permissive, allow child to make independent decisions *versus* parents appear dominating, over-protective toward child.
36	·22	Negavisitic, stubborn, disobedient, argumentative *versus* co-operative, compliant, obedient.

One encounters here the usual difficulty in separating C, ego strength, from G, super ego strength (since both typically share perseverance), on the one hand, and 0 (–) factor on the other (since C (–) and 0 typically share 'worrying, anxious'). Persevering, calm, stable and co-operative (lowest) are taken as the best definition of C, and no other pattern here has these, though Factor 4, D, below has some claim, and the closeness to 0 (–) Factor 12, is mitigated only by the latter's emphases on specifically fearful reactions, and its absence of general emotionality-vs.-calmness. If the reader sets this beside the C pattern as given in the Child Personality Questionnaire (Porter and Cattell, 1960, p. 26) however, he will immediately see the closeness of match.

TABLE 7

FACTOR 4—DEPRIVATION *versus* EXCITEMENT, D.

Item No.	Factor Pattern	Trait Description.
32	·64	Parents appear relatively permissive, allow child to make independent decisions *versus* parents appear dominating, over protective toward child.
31	·51	Parents appear generally accepting toward child *versus* parents appear generally rejecting toward child.
24	·42	Lacks nervous habits *versus* prone to nervous habits (e.g., thumb-sucking, nail-biting, pulling and twisting hair, grimacing)
14	·37	Calm, relaxed *versus* overactive, excitable, perhaps irritable.
38	—·25	Adventurous, bold, willing to take the chance of possible rejection or injury *versus* retiring, cautious
18	·19	Lethargic, pathetic, easily tired or fatigued *versus* alert, wide-awake, energetic.
12	·18	Lacking in self-confidence, easily discouraged or defeated *versus* confident (perhaps over-confident) of own ability and ideas.
13	·15	Placid, free from distress *versus* fearful, worrying, anxious.

Normally the central feature of D is excitability, with assertive attention-getting. The mixture of excitable with energetic, confident, yet nervous and fearful comes nearest here to the usual D. Also the tendency to pervade the whole personality (only 62 per cent. in hyperplane) more than other factors (except F and G) agrees. The new variables here—dominating and rejecting parent—raises crucial issues for further research on its nature.

TABLE 8

FACTOR 5—SUBMISSIVENESS *versus* DOMINANCE, E.

Item No.	Factor Pattern	Trait Description.
39	·62	Female *versus* male.
1	·45	Non-aggressive, kind, considerate *versus* aggressive, tends toward fighting, bullying, teasing, cruelty.
14	·38	Calm, relaxed *versus* overactive, excitable, perhaps irritable.
17	·33	Inattentive, absent-minded, prone to day dreaming, shows poor concentration *versus* very attentive to class proceedings.
4	·30	Quiet *versus* noisy, distracting in class.
8	− ·30	Dependent on teacher *versus* self-sufficient, independent.
22	·25	Practical minded *versus* imaginative.
21	− ·22	Polished in manner *versus* socially awkward, and clumsy.

The only factor with which dominance is occasionally confused are D, Excitability, above, and I (–), Harria (Table 12) below. Both E and I (–) are substantially sex (male) correlated in every group yet studied. The association of aggressiveness, activity, self sufficiency and ' socially clumsy and awkward ' (' unconventional ' in adults) suggest this is dominance (while ' esthetic ' marks I as different).

TABLE 9

FACTOR 6—DESURGENCY *versus* SURGENCY, F.

Item No.	Factor Pattern	Trait Description.
2	·56	Unpopular, generally disliked by other children *versus* popular, generally liked by other children.
15	− ·56	Depressed *versus* cheerful.
34	·46	Associates mostly with children of opposite sex *versus* associates mostly with children of own sex.
18	·43	Lethargic, apathetic, easily tired or fatigued *versus* alert, wide-awake, energetic.
5	− ·39	Prone to jealousy *versus* free of jealousy, feels accepted.
11	− ·20	Easily upset, overwhelmed by teasing of other children, yields easily to persuasion *versus* well poised, tough, sticks up for own rights when threatened.

The only problem usually met in recognizing the F pattern is its separation from H, since both involve active social behavior. But there is no difficulty here since the markers cheerful and energetic so clearly predominate. The s index of this pattern with Coans' study F is at P < ·02 and with Coan's H at P < ·25.

TABLE 10

FACTOR 7—SUPEREGO STRENGTH, G.

Item No	Factor Pattern	Trait Description.
25	—·89	Untidy, careless with respect to appearance, of self and belongings *versus* neat, tidy, orderly.
19	—·82	Irresponsible, frivolous *versus* responsible.
27	—·68	Careless, destructive of property of others *versus* careful with property of others
20	·59	Quitting, fickle *versus* persevering, determined.
16	·56	Changeable in interests, attitudes, opinions *versus* stable in interests, attitudes, opin;ons.
21	·48	Socially awkward, and clumsy *versus* polished in manner.
17	·38	Inattentive, absent-minded, prone to day dreaming, shows poor concentration *versus* very attentive to class proceedings.
7	—·32	Untrustworthy, dishonest *versus* conscientious, trustworthy.

This has all the marks of the super ego pattern, G, and needs no discussion.

TABLE 11

FACTOR 8—THRECTIA *versus* PARMIA, H.

Item No.	Factor Pattern	Trait Description.
35	·80	Prefers solitary pursuits *versus* gregarious, prefers games involving many children.
38	·75	Retiring, cautious *versus* adventurous, bold, willing to take the chance of possible rejection or injury.
28	·61	Secretive, reserved *versus* expressive, frank.
33	·58	Shy, bashful, seclusive, aloof, remains fairly isolated from other children *versus* mixes outgoing, freely with other children.
34	—·49	Associates mostly with children of own sex *versus* associates mostly with children of opposite sex.
25	·48	Neat, tidy, orderly *versus* untidy, careless with respect to appearance of self and belongings·
18	·45	Lethargic, apathetic, easily tired or fatigued *versus* alert, wide-awake, energetic.
21	·43	Socially awkward, and clumsy *vrrsus* polished in manner.

This has all the typical marks of H and its characteristic large variance contribution (including even the " tendency to mix more freely with the opposite sex," usually clearer at later ages, in this factor). The s index with the average pattern in Cattell's earlier collation for the pattern is P < ·05 and with Coan's P < ·05.

TABLE 12

Factor 9—HARRIA *versus* PREMSIA, I.

Item No.	Factor Pattern	Trait Description
39	—·58	Male *versus* female
29	·47	Lacking in artistic feeling *versus* aesthetically sensitive, aesthetically fastidious.
10	·41	Prefers not to be noticed *versus* demanding of teacher's attention.
22	·40	Practical minded *versus* imaginative.
9	·40	Submissive, follows lead of other children *versus* self-assertive, tends to dominate other children.
28	·39	Secretive, reserved *versus* expressive, frank.
11	·35	Well poised, tough, sticks, up for own rights when threatened *versus* easily upset, overwhelmed by teasing of other children, yields easily to persuasion.
33	·35	Shy, bashful, seclusive, aloof, remains fairly isolated from other children *versus* outgoing, mixes freely with other children.
5	·34	Free of jealousy, feels accepted *versus* prone to jealousy.

Again, the match to the previous concepts and empirical patterns for the factor is excellent. Out of feminine, esthetically sensitive, imaginative, demanding, easily upset, dependent and jealous only dependent is here missing. This is the typical environmentally produced pattern called premsia (as an acronym for *protected emotional sensitivity*). Mascalinity-femininity has a disportionate loading because of its perfect reliability !

TABLE 13

FACTOR 10—COASTHENIA, J.

Item No.	Factor Pattern	Trait Description
3	·39	Poor general health, prone to absence by reason of illness, or physical complaint *versus* of generally good health
9	—·32	Self-assertive, tends to dominate other children *versus* submissive follows lead of other children.
2	—·28	Popular, generally liked by other children *versus* unpopular, generally disliked by other children.
17	·18	Inattentive, absent-minded, prone today dreaming, shows poor concentration *versus* very attentive to class proceedings.
30	—·16	Follows instructions easily and accurately *versus* has difficulty following instructions.
37	·16	Trustful of others, readily accepts solicitude of others as sincere *versus* suspicious of others, ungrateful, rejects affection or solicitude.
35	—·15	Gregarious, prefers games involving many children *versus* prefers solitary pursuits.
34	·14	Associates mostly with children of opposite sex *versus* associates mostly with children of own sex

This is a pattern of such low loadings that it is almost certain to have its 'salients' obscured by erratics. It has been tentatively identified as J, Coasthenia (see Porter and Cattell, 1960, p. 33) which is marked by individualism, unpopularity, and lack of zest; but the match is not good and a second, but still poor possibility is L, Protension. It must await more intensive representation of variables crucial to its clarification.

TABLE 14

FACTOR 11—PRAXERNIA *versus* AUTIA, M.

Item No.	Factor Pattern	Trait Description
29	·63	Lacking in artistic feeling *versus* aesthetically sensitive, aesthetically fastidious.
22	·56	Practical-minded *versus* imaginative
35	—·35	Gregarious, prefers games involving many children *versus* prefers solitary pursuits.
3	—·35	Of generally good health *versus* poor general health, prone to absence by reason of illness or physical complaints.
24	·34	Lacks nervous habits *versus* prone to nervous habits (e.g., thumb-sucking, nail-biting, pulling and twisting hair, grimacing).
6	·30	Learns slowly *versus* learns fast.
35	·26	Neat, tidy, orderly, *versus* untidy, careless with respect to appearance of self and belongings.
34	—·17	Associates mostly with children of own sex *versus* associates mostly with children of opposite sex.

'Imaginative' and 'esthetically fastidious' typically load both factor I, Premsia, and factor M, Autia. That this factor is M is indicated by I being already present (Table 12 above), and by the fact that this lacks the traits of over-protectiveness characteristic of premsia. At the same time it has the essential meaning of Autia, as 'autistically wrapt up in one's own ideas, and lacking in practicality' as shown by 'prefers solitary pursuits,' 'imaginative' (this can be all things to all raters, but in this context, it evidently has the meaning of a vivid internal life), and 'careless of appearance.'

While this pattern has the main markers for the O factor of guilt proneness—notably 'worrying,' 'easily discouraged,' 'easily upset,' and 'dependent'—it also has some at lower loadings which normally go with protension (L factor or 'paranoid tendency'), such as 'prone to jealousy,' 'physical complaints' and 'easily upset.' There are two possibilities: (*a*) that one more factor should have been extracted, so that *possibly* this factor would have split into a larger O and a smaller L; (*b*) that in younger children a common pattern exists out of which guilt proneness persists later, as a direct development, and the paranoid 'protension' develops as a special reaction formation modification. This is an interesting theme which could be handled by a concentration of factor analytic research in this area.

TABLE 15

FACTOR 12—GUILT DEPRESSION OR SUPEREGO PRONENESS, O.

Item No.	Factor Pattern	Trait Description
13	·80	Placid, free from distress *versus* fearful, worrying, anxious.
26	·65	Adaptable, flexible *versus* rigid, has difficulty adjusting to changes or new situations.
12	–·62	Confident (perhaps over-confident) of own ability and ideas *versus* lacking in self confidence, easily discouraged or defeated.
3	–·60	Of generally good health *versus* poor general health, prone to absence by reason of illness, or physical complaints.
10	·59	Prefers not to be noticed *versus* demanding of teacher's attention.
11	·56	Well posed, tough, sticks up for own rights when threatened *versus* easily upset, overwhelmed by teasing of other children, yields easily to persuasion.
8	·50	Self-sufficient, independent *versus* dependent on teacher.
5	·42	Free of jealousy, feels accepted *versus* prone to jealousy
24	·38	Lacks nervous habits, *versus* prone to nervous habits (e g., thumb-sucking, nail-biting, pulling and twisting hair, grimacing)

IV.—SUMMARY AND CONCLUSION.

A factor analysis of checked personality rating data, on a standard and comprehensive set of variables, making systematic comparisons possible with researches on other samples and ages, has been described. The study is presented partly to show how more intensive technical attention to checks on (*a*) the number of factors to extract, and (*b*) the significance of the achieved simple structure, can produce closer agreement of factor analytic studies than has often prevailed, and can contribute basically to our concepts of child personality.

For the present research reaches a higher and virtually unimprovable hyperplane count compared with the alternative published analysis of the same data. Except for discrepancies in the two points named, the alternative study stands at a high technical level among experiments in this field and ends in factor patterns seven out of ten of which match seven present here. However, the present study gives very close matches indeed, in every factor (except possibly for J), to the standard list, indexed A through O, representing the collation of previous researches (Cattell, 1957). These twelve factors fit especially those found with children, by Coan (1957), Peterson and Cattell (1959), Huffman (1957), Koch (1942) and others, but also have clear continuity with twelve of the sixteen most fully described adult factors (Cattell, 1946; Degan, 1952; Huffman, 1957; Lorr, 1953; Norman, 1959; Thurstone, 1934; Types, 1958; and Wittenborn, 1951).

Unfortunately, the use of rating factors in experimental studies to evaluate maturationally and environmentally determined changes in individual children suffers from severe practical difficulties, due to the incomparability of different raters and observation situations. It is, therefore, particularly encouraging to find here that the behavior rating factors agree in number and meaning with those in a practical measuring instrument—the Child Personality Questionnare (Porter, 1960)—derived from similar factoring of *questionnaire* descriptions of behavior. Each of these factors is also found, though with two additions appearing with increasing age, in the High School Personality Questionnaire (Cattell and Beloff, 1960). However, as the present writer has pointed out, with more experimental and conceptual background, elsewhere (1961), this does not necessarily imply that the correlation should be unity between corresponding questionnaire and rating factors, because of the intrusion of instrument factors into the variance of both rating and questionnaire variables.

In view of the discrepancies which exist between the Digman study and the present one, in the meaning of a few smaller factors, it is most desirable that further independent studies with adequate rating conditions, *on the same variables*, should be carried out at this 6–8 year old level. Only then will it be timely to debate and experiment upon the hypotheses recently put forward (Cattell, 1957 ; Digman, 1963) for these behavior patterns.

However, in thus critically focussing on lesser discrepancies, one must not overlook the main conclusion which is one of agreement and, indeed, in an agreement across the appreciable cultural gap from the present writer's Caucasian subjects in the Mid-West and Digman's more than 50 per cent. racially Mongolian subjects in Hawaii. The recent cross cultural comparisons of independent simple structure factor analyses of the Sixteen Personality Factor Questionnaire (Cattell and Stice, 1962) in France (Cattell, Pichot and Rennes, 1961), Italy, Japan and the U.S.A. show clearly that such basic source traits as cyclothymia-sizothymia, dominance, surgency, ego strength, etc., described in ratings above, persist in essential form in all cultures investigated. It is part of the special value of the Digman data that it contributes towards checking, in the rating field, this theory that the A through O factors are common to humanity.

V.—REFERENCES.

BARGMANN, R. (1954). Signifikanz untersuckungen der Einfach en Struktur. *Mitt, fur Mat Stat*, Sonderdruck, Wurzbarg : Physica-Verlag.

BURT, C. L. (1940). *Factors of the Mind*. London : University of London Press

CATTELL, R. B. (1945). The description of personality *Amer. J. Psychol*, 58, 69-90.

CATTELL, R. B. (1946). *The Description and Measurement of Personality*. New York : Yonkers, World Book.

CATTELL, R. B (1957). *Personality and Motivation Structure and Measurement*. New York : Harcourt, Brace and World

CATTELL, R. B (1961). Theory of situational, instrument second order and refraction factors in personality structure. *Psychol. Bull.*, 58, 160-174.

CATTELL, R. B., and BELOFF, H. (1960). *The High School Personality Questionaire National Found. for Educ. Res.* London : Wimpole Street.

CATTELL, R. B., and COAN, R A (1957). Child personality structure as revealed in teachers' ratings *J. Clin Psychol*, 13, 315-327.

CATTELL, R B., and DICKMAN, K. (1962). A dynamic model of physical influences demonstrating the necessity of oblique simple structure *Psychol Bull.*, 59, 389-400

CATTELL, R. B., and MUERLE, J. L. (1960). The ' maxplane ' program for factor rotation to oblique simple structure. *Educ. Psuchol Measmt.*, 20, 569-590.

CATTELL, R. B., PICHOT, P., and RENNES, P. (1961) Constance interculturelle des facteurs de personnalité mesures par le test 16 P.F. II.—Comparaison franco-américaine. *Revue de Psychologie Appliquée*, 11, 165-196

CATTELL, R. B., and PAWLIK, K. (1963). A new test of significance of simple structure. (Laboratory Rept, Univ. of Illinois, Lab. Person. Assess.)

CATTELL, R. B., and STICE, G. F. (1962) *The Sixteen Personality Factor Questionnaire.* N.F.E.R., London.

DEGAN, J. W. (1952). *Dimensions of Functional Psychosis*. Richmond, Va. : Byrd Press.

DIGMAN, J. (1963). The principal dimensions of child personality as inferred from teachers' judgments. Child development 34, 43-60.

GORSUCH, R. L. (1963). A culturally sampled check on the second order structure among questionnaire personality primaries. (In Press).

HUFFMAN, P. W. (1957). See Cattell, R. B., ref. 1957, above

KOCH, H. L (1942) A factor analysis of some measures of the behavior of pre-school children... *J. Gen. Psychol*, 27, 257-287.

LORR, M., and RUBENSTEIN, E., and JENKINS, R. L. (1953). A factor analysis of personality ratings of out-patents in psychotherapy. *J. of Abnorm. Soc. Psycgol.*, 48, 511-514.

NORMAN, WARREN T. (1959). A check on behavior rating factors in adults. *J. Abnorm. and Soc Psychol.* In Press.

PASCAL, G. R., and JENKINS, W. O. (1955). A study of early environment and its relation to adult behavior. *J. of Abnorm. Soc. Psychol.*, 43, 279-284.

PETERSON, D. R., and CATTELL, R. B. (1959). Personality factors in nursery school children as derived from teachers' ratings. *J. Consult. Psychol.*, 23, 562.

PINZKA, C., and SAUNDERS, P. R. (1954). Analytic rotation to simple structure : II.—Extension to oblique solution. (RB-54-31) *Educational Testing Service Bulletin* (Multilithed)

PORTER, R., and CATTELL, R. B. (1960). *The Child Personality Questionnaire.* Champaign, Ill. : Institute of Person and Abil. Testing, 1602, Coronado Drive.

SELLS, S B (1961) (Edit.). The dimensions of stimulus situations which account for behavior variance. Symposium No. 1, Psychology Department, Texas Christian University, Texas.

SOKAL, R. R. (1959). A comparison of five tests for completeness of factor extraction For *Kansas Acad. Sci.*, 62, 141-1952

THURSTONE, L. L. (1934) The vectors of the mind. *Psychol. Rev.*, 41,1-35.

TSUJIOKA, B. Similarities and differences in personality structure in Japan and the U.S.A., as shown by factoring the 16 P.F. (In Press)

TUCKER, L., in Thurstone, L. L (1938). *Primary Mental Abilities.* Chicago : Univ. of Chicago Press

TUPES, ERNEST, C., and CRISTAL, RAYMOND, C. (1958) Stability of personality-trait rating factors obtained under diverse conditions. *U.S.A.F., W.A.C:D. Tech. Note*, No. 58-61.

WITTENBORN, J. R (1951). Sympton patterns in a group of mental hospital patients. *J. of Consult. Psychol.*, 15, 290-302.

(*Manuscript received 1st January*, 1963)

AN EXPERIMENT ON THE EFFECTS OF MENTAL SET IN THE BROOK REACTION TEST

By D. H. HARGREAVES, A. W. HEIM and K. P. WATTS.
(*The Psychological Laboratory, Cambridge*)

SUMMARY. In the Brook Reaction test an attempt is made to assess the subject's interests. Words potentially denoting one of six selected interests were, in this experiment, presented consecutively to an experimental group, and in randomised order to a control group. Some degree of mental set was induced in the experimental subjects, reaching significance level in the case of one interest only.

I.—INTRODUCTION.

IN a preliminary note on the Brook Reaction test, Brook and Heim (1960) confirmed the hypothesis that people's individual interests may be revealed by their responses to an ambiguous word association test. They used 106 orally presented stimulus-words each of which had two or more meanings and at least one of these meanings related to one of several specific interests. The subjects were thirty Cambridge university students ; the non-test criterion consisted of interview ratings ; and significant (tau) correlations between the two criteria were obtained for seven of nine interests.

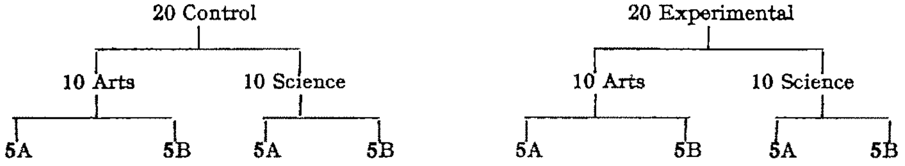
In the original experiment, the stimulus-words were presented in random order, words potentially denoting similar interests being generally separated from one another. Half of the subjects took the test words in the order 1–106 (order A) and the other half took the test words in the reverse order (B). No significant differences were found as a result of these two orders.

It was thought interesting, however, to see whether a mental set could be deliberately induced in the subject, by presenting several similar interest words consecutively ; to ascertain how many successive words of similar interest-category would be necessary to effect such a set ; and to determine the extent of individual variation with regard to the set. We were interested too in possible interest-differences between arts men and scientists. With these aims in view, the present experiment was designed.

II.—EXPERIMENTAL DESIGN.

The subjects consisted of forty Cambridge university undergraduates. Twenty were arts men, reading English, law, modern languages, theology or history, and twenty were scientists, reading natural or mechanical sciences or mathematics. These forty were divided into two groups, as follows :

TABLE 1



Thus, both the Control and the Experimental groups contained ten arts men and ten scientists, half of whom took the words in one order and the other half, in the reverse order. All the subjects were male and aged 18–23 years.

They were told that the experiment was on reaction times and were asked to give the first word or phrase that entered their mind immediately after hearing each test-word. A stop-watch was in evidence. As in the previous experiment, the ambiguity of the test-words was not recognised by the majority of the subjects.

Since the original experiment the number of interest-categories had been increased from nine to eighteen. For the Experimental group, six of these were chosen: *entertainment*, *food and drink*, *intellectual*, *practical*, *religion* and *sport*. The stimulus-words were then grouped, classifying any that could be so grouped under one of these six heads. This left a total of thirteen ' neutral ' words that is words connoting for instance, *aesthetic* or *political* or *business* interests—categories which differed from the six interests chosen for the ' mental set ' investigation. Six (or seven) of these neutral words occurred at the beginning, and seven (or six) occurred at the end of the test, according to whether order A or order B was used. For the Experimental A subjects, the order was as follows:

neutral words (e.g., ' labour ')
  entertainment words (e.g., ' cast ')
a linking word (e.g. ' character ')
  intellectual words (e.g., ' books ')
a linking word (e.g., ' style ')
  sport words (e.g., ' match ')
a linking word (e.g., ' pole ')
  practical words (e.g., ' construction ')
a linking word (e.g., ' flour ' (or ' flower ')
  food and drink words (e.g., ' bacon ')
a linking word (e.g., ' service ')
  religious words (e.g., ' minister ')
neutral words (e.g., ' sale ' (or ' sail ').

The B order was the exact reverse. It may be seen that the linking words have possible associations with the preceding and succeeding ' set ' groups of words. For example, the stimulus-word ' service ' occurs between the *food and drink* and the *religion* groups. The response ' restaurant ' would be taken as a *food and drink* association whilst the response ' church ' would be taken as a *religious* association. Similarly, with the linking word ' character ': the response ' play' would be classed as *entertainment* whereas the response ' mind ' is classed as *intellectual*.

After having taken the test, each subject completed a seven-point rating scale on himself with regard to the relevant interests. In all but five cases a second interests-rating was obtained on each subject, by a person who knew him well. Good agreement was found between the pairs of ratings, a difference of more than one point being extremely rare.

III.—RESULTS.

A—The grouping together of stimulus-words with similar interests.

TABLE 2

Interest Category	Frequency		χ^2	p	Mean Rating	
	Control	Exptal.			Control	Exptal.
Entertainment	57	71	0·20	0·65	3·9	3·2
Food and drink	112	156	4·01	0·045	3·1	3·1
Intellectual	38	54	0·11	0·74	3·5	3·5
Practical	80	85	0	1·0	2·0	1·9
Religious	118	130	0	1·0	2·8	3·1
Sport	98	154	2·13	0·14	2·6	2·7

Table 2 shows the results of our attempt to induce a mental set, by arranging the stimulus-words in six concentrated groups. The second and third columns show the frequency of interest responses, given by the Control and the Experimental subjects, respectively. It may be seen that in every case, the frequency of the Experimental subjects exceeds that of the Control subjects. This difference always is in the expected direction, but in only one case, that of *food and drink*, is the difference statistically significant. This ' set ' group included such words as ' roll ' ('role'), ' date,' ' mint,' and ' palate ' (' palette '). The values of χ^2 are given in column 4 and those of p, in column 5.

The pair of columns on the extreme right show the mean *rating* on each of the ' set ' interest categories for, respectively, the Control and the Experimental subjects. It is thought worthwhile to give this since it was just possible that the Control/Experimental differences in columns 2 and 3 were due to genuine differences in the interests of the two groups of subjects. The ratings make it clear, however, that the finding is, in fact, a result of the induced ' set,' the biggest rating difference being *in favour of the Control group* (*entertainment*) as opposed to the Experimental group, and two of the other interest-categories (*intellectual* and *food*) yielding identical means. None of the differences between mean ratings approach significance level.

A contingency table for all the ' set ' groups of stimulus-words taken together (twelve degrees of freedom) yielded a χ^2 of 20·504 ; $p=0{\cdot}058$. This borderline significance level of confidence for the Experimental testing as a whole, suggests that there is an overall effect of ' set,' although the effect was insignificant for five of the six categories considered separately.

B—The linking words.

It was hypothesised that if the A *versus* B presentation of the stimulus-words had any effect, this would appear in the responses to the linking-words : that the linking word—and perhaps one or two of the stimulus-words on either side of it (which were also carefully chosen)—might show the effect of the ' set ' group that was taken immediately before. In seven cases out of ten this tendency was found and in one case only was the opposite tendency clearly shown.

An example may help to clarify this. The linking word ' flour ' (or ' flower ') occurs between the *practical* and *food and drink* ' set ' groups. ' Flour ' is flanked by ' pick-up ' and ' joint,' these two words being, respectively, the last of the *practical* stimulus-words and the first of the *food and drink* stimulus words in order A. Subjects taking the test in this order produced from these three stimulus words a total of seven *practical* responses and two *food and drink* responses. Those taking the test in order B produced only three *practical* responses but eighteen *food and drink* responses.

C—Arts/Science differences.

The differences between the arts men and the scientists on (i) the test and (ii) the ratings, were inspected. For this purpose, both the Control and the Experimental subjects were included, since the groups were equal in number and contained equal numbers of arts men and scientists. The four interests which yielded the greatest differences were *practical, scientific, clothing* and *people*. The scientists displayed greater interest in the former two, and the arts men in the latter two. In all four cases, the test scores and the ratings agreed with one another—very strikingly in the case of the *practical* and *scientific* categories.

IV.—DISCUSSION.

This is a small-scale experiment designed primarily to see whether 'set' is likely to arise, and to have a deleterious effect, in the Brook Reaction test as normally given, that is, with the word order randomised with relation to potential interest value. From this point of view, the experimental results are satisfactory: when words connoting similar interests were given consecutively, for six interest-categories, the effect of 'set' reached significance level in only one instance. It is true that for the other five interests, some 'set' effects were discernible, but these were far from significant. When the results are summated, the influence of 'set' on the Experimental group is of borderline significance only. Furthermore, at least three or four consecutive words of potentially similar interest value were necessary before a 'set' was effected. The danger, therefore, of inadvertently introducing a 'set' in the Brook Reaction test proper (as taken by the Control group), which might obscure the relationship between stimulus-words and genuine interests is negligible.

There was a certain amount of individual variation among subjects, both in the Experimental and the Control group. For example, one subject in the Experimental group gave the identical response to each of the following (consecutive) stimulus-words: 'box,' 'race,' 'bowl.' The response in each case was 'horse' and the subject commented later than he had been unable to rid his mind of the idea 'horse'! These three stimulus-words appeared towards the middle of the *sport* 'set' group for the Experimental subjects.

On the other hand, a Control subject gave four out of five *science* responses to what purported to be a mixed group of five stimulus-words. The stimuli and responses were as follows:

Stimulus-word	*Response*
construction	diagram
punch	drunk
arc	circle
joint	Fisher (geology)
plot	curve

Such findings were rare, however, and probably tell us more about the individual subject than they do about the test. The results in general suggest that (i) a misleading mental set is unlikely to be inadvertently introduced, when the words are presented as to the Control subjects, in random order and (ii) that interest-differences which obtain between arts men and scientists, as such, are reflected at least as strongly in the test score as they are in the ratings.

ACKNOWLEDGEMENTS.—The major part of this experiment was carried out by D. H. Hargreaves as part of his research project for the Part II Psychology Tripos. The other two authors wish to express their thanks to the Medical Research Council for financial support.

V.—REFERENCE.

BROOK, D. F., and HEIM, A. W. (1960). A preliminary note on the Brook Reaction test. *Brit. J. Psychol.*, 51, 347-356.

(*Manuscript received* 12*th December*, 1962.)

THE DEVELOPMENT OF SCIENTIFIC CONCEPTS IN CHILDREN (2)

By W. H. KING

(*University of London Institute of Education*)

SUMMARY. This paper describes the responses of children to twenty science questions and relates these responses to age and ability represented by scores on a verbal and non-verbal test. A relationship is established for certain types of question between their facility values and the grades obtained by the children on the criteria. Rate of growth of scientific knowledge is found to be similar for boys and girls ; boys and girls of 10 years of age show no significant differences in knowledge of science concepts but there is significant differentiation in secondary school children.

I.—INTRODUCTION.

A PREVIOUS paper in this *Journal* (Vol. XXXI, p. 1, 1961) discussed the results of an investigation dealing with the growth of children's scientific knowledge with increase of chronological age. The replies to seventy questions were examined and eighteen of these showed quite clearly a steady increase of correct answers with increase in age (*loc. cit.*, Appendix 2). These eighteen questions with two more drawn from the original list were used as a basis for further study. There was no evidence from the previous research that any of these questions were unduly favourable to either sex. The form of presentation of these questions was decided at a discussion with the group of teachers who had volunteered to carry out this research in their schools.

II.—THE INVESTIGATION.

(1) *Material.* The twenty questions in the Science Schedule were produced on a standard form on which the replies could be recorded objectively by the children. All the children were given the Simplex Junior Intelligence Scale and the Raven Progressive Matrices (Revised Order, 1956) so that for each child the following information was available :

(*a*) chronological age,
(*b*) raw score on verbal test,
(*c*) raw score on non-verbal test,
and (*d*) responses to each of the twenty questions of the Science Schedule.

(2) *Subjects.* Teachers in twenty-six schools offered to co-operate in this research. This would have involved the testing of approximately 4,800 children. To reduce the size of the experiment to manageable proportions, the complete battery of test material was given to every fifth or sixth child (depending on size of class) on each class register. By this arrangement, 801 children (368 boys and 433 girls) were tested ; they formed a random group from the participating schools even though the schools themselves were not randomly selected. The oldest was 17 years 3 months and the youngest 5 years 4 months. This wide age-range was a decisive factor in the selection of suitable published tests. Of the twenty-six schools, seventeen were mixed primary schools, one infants' school, two secondary modern girls' schools, two secondary modern boys' schools, one mixed secondary modern school, one girls' grammar school, one girls' comprehensive school and one mixed independent school. Twenty of these schools were in the Greater London area and six in the Home Counties ; 528 children were at the primary stage and 273 at the secondary stage.

Table 1 shows the age distribution according to sex and type of school.

TABLE 1

Age	17		16		15		14		13		12		11		10		9		8		7		6		5		Boys	Girls
Sex	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G	B	G		
Primary													46	54	79	71	62	54	46	46	28	15	12	8	3	4	276	252
Sec. Modern ..							11	29	22	38	21	33	3	9													57	109
Sec. Grammar	1	0	3	1	4	10	9	19	8	14	5	19	5	9													35	72
	1	0	3	1	4	10	20	48	30	52	26	52	54	72	79	71	62	54	46	46	28	15	12	8	3	4	368	433

The seventeen children in the infants' school and the eleven boys and the eleven girls in the one independent school who were under 11 years old were grouped under the primary heading. The forty-two girls in the comprehensive school were following secondary modern courses and were, therefore, included in the secondary modern figures.

The smaller number of boys in the grammar school group (thirty-five boys compared with seventy-two girls) made the sample less representative than was desired.

Table 2 summarises the nature of the sample by sex and stage of education.

TABLE 2

	Boys	Girls	Total
Primary	276	252	528
Secondary ..	92	181	273
Total	368	433	801

III.—THE RESULTS.

The means and standard deviations on the four variables for all the children and for boys and girls taken separately are given in Table 3.

TABLE 3

	Mean Scores				Standard Deviations			
	Age	Simplex (100)	Matrices (60)	Science Schedule	Age in months	Simplex	Matrices	Science Schedule
Boys ..	10 : 6·2	53·5	35·5	11·95	26·1	21·6	11·5	3·3
Girls ..	11 : 2·3	59·7	38·5	11·50	26·8	20·9	12·0	3·2
Total ..	10 : 10·6	57·1	37·1	11·71	26·8	21·4	11·9	3·3

The average age of the girls exceeded that of the boys by eight months and this difference was significant at the ·01 level. The difference in the means of the Simplex scores and Matrix scores were significant (·01 level). On the Science Schedule, considered as a test, the boys' scores were just significantly higher (C.R.=1·96).

The intercorrelations of the four variables were calculated separately for boys and girls and since there were no significant differences, Fisher's Z-technique was used to calculate the correlations on the complete 801 sets of results which are recorded in Table 4.

TABLE 4

	Simplex	Matrices	Science Concepts
Age..........	0·682	0·570	0·504
Simplex......	—	0·767	0·677
Matrices	—	—	0·573

These high correlations might be expected because of the wide range of age and ability of the children. The Simplex Junior is the best predictor of total Science Scores and is significantly better (·01 level) than either chronological age or Matrices; Matrices in turn are a better predictor than age (significant at ·05 level).

The largest group of children tested in a single year group was those whose ages were in the range 10 to 10: 11; in all there were seventy-six boys and sixty-six girls from seventeen junior schools. The three boys and five girls from the independent school were not included. The separate means and standard deviations for these boys and girls (Table 5) on the Matrices, Simplex Junior and the Science Schedule did not differ significantly, neither did the intercorrelation of these variables (Table 6).

TABLE 5

MEANS AND STANDARD DEVIATIONS OF THE THREE VARIABLES FOR CHILDREN, AGES 10 : 0 TO 10 : 11.

	Mean Scores			Standard Deviations		
	Simplex	Matrices	Science Schedule	Simplex	Matrices	Science Schedule
Boys	58·1	38·0	12·2	13·3	9·0	2·4
Girls	62·2	40·6	11·6	14·1	9·3	2·6

TABLE 6

INTERCORRELATION OF THREE VARIABLES FOR CHILDREN AGES 10 TO 10 : 11

		Science Schedule	Simplex
Matrices .. .	Boys	0·359	0·650
	Girls	0·314	0·634
	Total	0·34	0·64
Science Schedule ..	Boys	—	0·399
	Girls	—	0·562
	Total	—	0·52

Comparing these results with those obtained from the extended age-range, the fall in values of the intercorrelations is to be expected ; although the 10-year-old girls still obtain higher mean scores on the Matrices and Simplex Junior and the boys higher mean scores on the Science Schedule, these differences are no longer significant. This suggests that had there been a more representative sample of grammar school children, there might have been a narrowing of the difference in means on the Matrices, but a widening of the mean scores on the Science Schedule.

Dr. Gertrude Keir reported correlations of 0·538 for boys and 0·588 for girls between the Simplex Junior and Matrices for a population of primary and secondary modern children using the 1938 version of the Progressive Matrices.[1] The higher correlations now obtained in the present study may be due to the improved version of this test.

With age partialled out, there were no significant differences between boys' and girls' correlations on the three remaining variables. For the pooled scores, the partial correlation between Simplex Junior and the Matrices was 0·629, compared with 0·64 for a single year group ; similar partial correlations between Matrices and the Science Schedule were 0·403 for the pooled scores and 0·34 for the single year group.

Allowing for the fact that the scores in the single year group are still influenced by age, the corresponding correlations of the scores of the children in the complete age-range, when age is partialled out, do not differ unduly from the correlations obtained in the 10 to 10 : 11 group. The correlation between the Simplex Junior and the Science Schedule, with age partialled out, was ·527. Again, the verbal test was a better predictor of science scores (significant at ·05 level). Further partialling out of variables yields figures which are less reliable but the results are recorded as they show the relative influences of verbal and non-verbal scores on the Science scores.

TABLE 7

Variables partialled out	Correlations between	Boys	Girls
Age, Simplex	Matrices and Science scores	0·174	0·100
Age, Matrices	Simplex and Science scores	0·320	0·412

The boys' correlations differ significantly (·05 level) and the girls' correlations also (·01) but the differences in the boys' and girls' correlations are not significant.

As a test, the Science Schedule has a reliability of 0·604.[2] It was slightly more reliable for boys (0·60) than for girls (0·55). The corresponding reliabilities of the Simplex Junior and the Matrices, obtained by the same method were 0·96 and 0·91* respectively.

The lowest correlation (0·504) is between age and science scores. The effect of age on the science scores was shown graphically in the previous research (*loc. cit.*).

* *loc. cit.*, p. 143. In a discussion of the reliability of the Matrices (1938), Keir reports a reliability coefficient of 0·76 (split half method) and 0·64 for successive trials on a small sample of children.

The differences in the mean scores of the two sexes on the Simplex Junior were significantly different for the full age-range but not for the 10 to 10 : 11 children. It has been shown that the girls did significantly better than the boys on the Matrices for the complete age-range (·01 level) but this difference was not significant for the 10-year-old group. This may be evidence again of the lack of a larger grammar school boy representation in the sample and if this is so in a fully representative sample, the boys' scores on the Science Schedule would have been more than just significantly higher (C.R. = 1·96) than the girls : nevertheless, there was no significant sex difference in the science, matrix or verbal test scores (means and standard deviations) for the 10-year-old group. Such differences as have been found in the sample of older children may be the results of differing types of schooling and a further study of the groups of children in different types of school in the age-range 13 to 15 years might supply a more positive answer.

The multiple regression equation to predict science scores confirmed the influences of the other three variables—verbal, non-verbal and age. This predictive equation in score form is

$$X_{\text{Science}} = 0{\cdot}82\ X_{\text{Simplex}} + 0{\cdot}034\ X_{\text{Matrices}} + 0{\cdot}009\ X_{\text{age in months}} + 4{\cdot}67$$

giving a multiple correlation coefficient of 0·684. Approximately 47 per cent. of the variance of the Science test is attributable to the joint action of age, verbal and non-verbal variables ; 36 per cent. and 7 per cent. are the contributions of the verbal and non-verbal tests and 3·5 per cent. to the differences in age.

There were too few children in the 5-year-old, 16- and 17-year-old groups so these children were included in the 6-year group and 15-year groups, respectively. Twenty graphs of the percentage correct response for ages 6 to 15 were superimposed on the corresponding graphs for the same twenty graphs obtained in the previous study of primary school children. The two sets of graphs were similar for both primary school children ; such differences as existed might be attributed to : (*a*) differences in sampling, (*b*) the standardising of the form of the answers in the second study and (*c*) fewer questions in each section which cut down the chances of practice.

So far, the total scores of the twenty science questions have been compared with age, Matrix scores and verbal reasoning scores for boys and girls. An analysis by age, verbal and non-verbal score was obtained for each of the science questions. The raw scores were divided into five grades determined by their 20th, 40th, 60th and 80th percentiles, separately for boys and girls. The corresponding end scores, were calculated, checked graphically and plotted against correct percentage response for each science question. The graphs illustrated the similarity of trends in boys' and girls' responses. There were slight differences between boys and girls that might be attributed partly to sampling errors, but every graph showed a steady increase of correct percentage response throughout the grades E to A of the verbal and non-verbal tests.

Direct comparisons were possible between percentage correct replies on each of the science questions and the ages and grades on the Matrices and Simplex Junior tests. From the eighty graphs comparisons were made between question difficulty (as measured by percentage correct replies) and age, verbal grade and non-verbal grade. For example, in the science question (No. 15. Are all things that move living?) at least 60 per cent. of the answers were correct at a minimum age of 9 years ; also with this 60 per cent. facility value, boys and girls had to be graded D at least on the Simplex Junior, and on the Matrices, boys required a C grading and the girls a D grading at least. The choice of a facility value of 60 per cent. was quite arbitrary, but seemed a reasonable figure to take for purposes of comparison. The ages and grades are given in Table 8.

TABLE 8

PATTERN OF GRADES ON VERBAL AND NON-VERBAL GRADES WITH AGE CORRESPONDING TO 60 PER CENT. FACILITY VALUE ON EACH SCIENCE ITEM.

	Science Quest'ns	2	7	14	8	9	6	3	15	11	4	12	16	13	10	5	1	18	7	19	2
Age ..	Boys & Girls ..	5/6	5/6	5/6	7	8	8	8	9	9	9	9	9	11	14	14	14	15	—	—	—
Verbal Grades	Boys ..	E	E	E	E	D	E	D	D	D	D	D	B	C	C	—	—	—	—	—	—
	Girls ..	E	E	E	E	D	D	D	D	C	C	B	B	B	—	—	—	—	—	—	—
Non-Verbal Grades	Boys ..	E	E	E	E	E	E	E	C	D	C	C	B	B	C	—	—	—	—	—	—
	Girls ..	E	E	E	E	E	D	D	D	C	C	C	B	A	—	—	—	—	—	—	—

Questions 1, 5, 7, 18, 19 and 20 were too difficult and the only guide for this level of accuracy that can be deduced is that for Questions 1, 5 and 18 the age should be at least the age recorded, i.e., 14, 14 and 15, respectively. Less than 60 per cent. of the children, at whatever their age or grade, gave correct replies to Questions 17, 19 and 20. These questions were too difficult. A lowering of the facility value will reduce corresponding ages and verbal and non-verbal grades for each science concept.

Question 13 dealing with the relative speed of a train of cogs, brings out the differences in the replies of boys and girls. To obtain 60 per cent. correct answers to this question, the children should be at least 11 years old; the boys are likely to be a point above the fortieth percentile in the verbal test, girls above the sixtieth percentile; and the boys over the sixtieth percentile in the non-verbal test and the girls at least at the eightieth percentile.

IV.—DISCUSSION.

The correlation between the scores of the verbal and non-verbal test (0·767) was the highest correlation recorded. This may be due to the wide range of age and ability of the children as well as the common factorial element of the two tests. There was some evidence that the verbal test was a better predictor of success in the science questions than the non-verbal test. The difference of the correlation coefficients was highly significant statistically. Further evidence was provided by the relative proportions of the variance of the science test attributable to the verbal and non-verbal scores. The predictive powers of the three variables are in order (*a*) Simplex Junior test, (*b*) Raven's Progressive Matrices and (*c*) age. The science test was given in verbal form and it would seem, therefore, that a scholastic test with large g : v component might *a priori* be a better predictor of scientific attainment in the present study than a test which has a large 'g' loading.[3]

Boys were consistently better than girls on both the verbal and non-verbal criteria in questions relating to (*a*) shadow of post; (*b*) weights on see-saw; (*c*) rotation of cog-wheels; (*d*) shadows formed by edge of cardboard circle. These questions were not intrinsically more difficult than others where these differences were not apparent; for example, there were no differences in the pattern of responses to questions 18, 19 and 20 which candidates of whatever age and ability found difficult. These three questions demanded imagination of the shape of the cut surface but the other four types of question required the application of mechanical principles which for most children may not have formed part of formal science lessons. If this is so, it seems likely that boys tended to draw more on practical out of school experience.

The differences in the sample of boys and girls who took part in this study have already been noted. The girls were older, did better on the Matrices and Simplex Junior but the boys obtained significantly higher scores on the Science questions but with the 10 to 10 : 11 group there were no significant differences in any of the variables. It would appear that the boys' superiority on the science questions does not come about until after the children have left the primary school.

When the percentile grades were determined separately for boys and girls with the intention of equating the two distributions, the graphs of percentage correct response, first with the verbal grades and then with the non-verbal grades, gave graphs which were nearly identical so that allowing for differences in the boys' and girls' samples, the rate of growth of correct science replies compared with achievement on the verbal and non-verbal tests were almost identical. From this it might be deduced tentatively that differences in the attainments of boys and girls in science, e.g., at ' O ' Level, may not be due so much to differences in verbal and non-verbal ability, but to other non-academic influences, e.g., teaching, interest motivation, self-involvement, etc., with which this study was not concerned.

ACKNOWLEDGEMENTS.—The assistance in discussion and supervision of testing by the following teachers is gratefully acknowledged: Miss B. S. Best, Miss E. M. Burton, Mrs. R. Conroy, Miss N. M. Harn, Miss N. Kenefick, Miss K. G. Lighten, Miss F. E. Marshall, Miss E. M. Pointon, Miss L. E. Whitehead, Miss J. B. Whyte, D. Billyard, F. G. Bird, C. E. Butler, R. Calvert, D. Casson, D. H. Cleife, H. C. Deverson, J. K. Draper, R. E. Gordon, S. G. Melhuish, J. M. Muir, J. T. Nye, M.A., D. Oakley, B. Pickles, E. W. Self, F. J. D. Taylor, C. V. S. Williams.

V.—REFERENCES.

[1] KEIR, GERTRUDE (1949). Progressive matrices applied to school children. *Brit. J. Psychol.* (*Stats.*), Vol. II.

[2] GARRATT, H. E. (1959). *Statistics in Psychology and Education*, p. 341. Longman.

[3] VERNON, P. E. (1950). *Structure of Human Abilities*, p. 72. University of London Press.

(*Manuscript received 17th December*, 1962.)

VI.—APPENDIX 1 (A).

SCIENCE SCHEDULE. PERCENTAGE OF CORRECT SCIENCE SCORES.

Junior Simplex Percentiles.		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
80	B	58	95	92	94	54	85	88	100	85	77	94	90	85	100	88	77	58	29	25	27
	G	52	96	94	94	49	89	90	96	84	55	91	81	66	95	88	77	18	60	27	25
60	B	46	100	84	87	47	87	83	99	88	72	87	87	70	93	87	70	51	26	17	17
	G	35	85	73	87	32	89	78	92	73	55	73	72	65	86	80	67	15	42	19	11
40	B	46	94	76	82	31	87	84	95	86	61	81	71	62	85	74	56	27	35	13	14
	G	31	81	72	69	43	78	77	91	88	48	66	52	52	79	68	59	7	33	15	1
20	B	34	80	71	61	33	79	73	89	77	49	66	69	40	81	60	59	17	26	10	1
	G	30	73	68	54	38	78	83	82	76	41	51	41	31	80	63	70	13	32	8	1
0	B	32	77	58	46	23	61	76	62	55	38	40	49	34	93	44	50	12	17	13	
	G	21	64	58	48	19	56	92	68	59	26	30	37	23	74	27	45	8	22	16	1

VII.—APPENDIX 1 (B).

SCIENCE SCHEDULE. PERCENTAGE OF CORRECT SCIENCE SCORES.

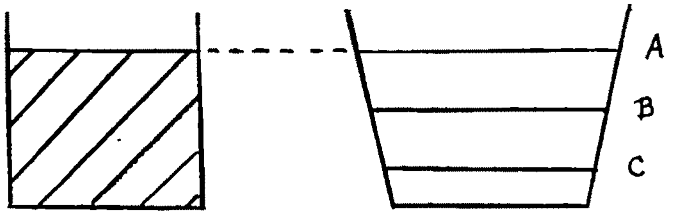
Matrices Percentiles		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	2
80	B	52	97	94	92	45	88	81	99	82	73	96	84	77	93	88	71	53	34	23	2
	G	56	96	92	92	50	90	90	95	83	58	84	80	77	95	92	73	23	58	36	3
60	B	51	100	80	83	36	83	83	96	86	71	87	88	61	90	81	74	37	31	20	2
	G	37	85	85	86	34	87	82	90	80	47	77	65	54	87	71	72	10	38	15	1
40	B	38	85	70	78	35	85	78	95	86	50	72	74	53	89	76	59	27	23	13	
	G	35	85	75	72	37	82	65	92	81	44	71	68	39	90	64	58	11	39	11	
20	B	42	81	66	59	35	75	76	85	68	47	60	56	44	87	53	46	17	29	11	
	G	29	71	70	55	44	79	79	85	80	38	56	43	44	79	63	65	11	31	8	1
0	B	24	82	63	51	27	68	77	66	66	42	46	49	30	86	45	55	18	16	11	1
	G	23	70	57	59	21	58	77	74	61	36	36	39	37	72	48	57	8	22	11	1

Percentiles were calculated from boys' and girls' combined distributions.

VIII.—Appendix 2.

Estimate the time in seconds between two taps. (1)

A jar is partly filled with water. A stone is gently dropped into the water. Is the level of the water now higher, lower, or the same as it was before ? (2)



Look at the diagrams

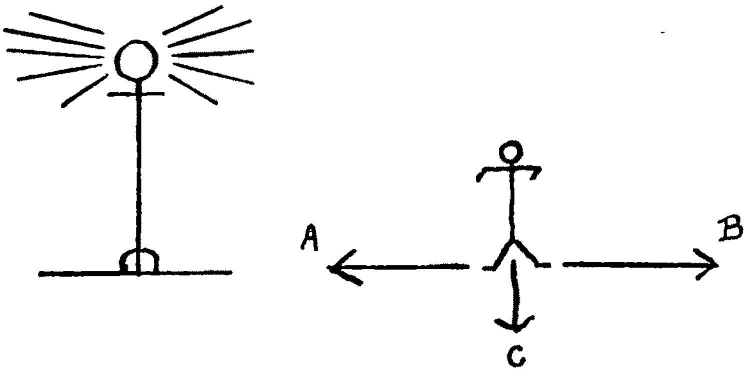
The first represents a glass of water. All the water is poured into the second glass. If the bottoms of both glasses are the same size, will the new level of the water be at A, B or C ? (3)

Imagine two pieces of plasticine which have the same size and weight. One lump is rolled into the shape of a pencil. Will the two pieces now have the

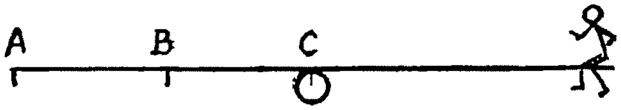
(*a*) same weight ? (yes, no ?) (4)
(*b*) same volume ? (yes, no ?) (5)

Ice turns into water, sugar dissolves in water.

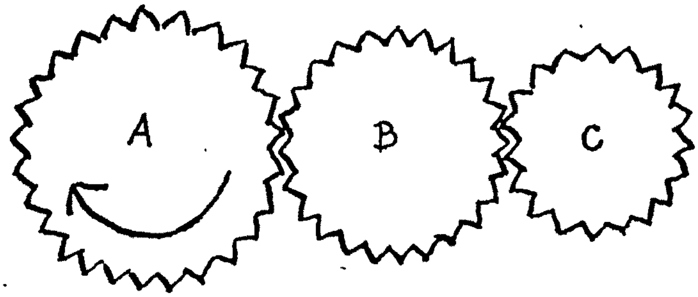
Does steam turn into water ? (yes, no ?) (6)
Does salt turn into water ? (yes, no ?) (7)
Does hail turn into water ? (yes, no ?) (8)
Do clouds turn into water ? (yes, no ?) (9)



The drawing shows a man standing in a road in the light of a lamp post. Will the man's shadow fall along the lines A, B or C ? (10)



Here is a diagram of a seesaw with a girl sitting at one end (*c*).
If a boy is twice as heavy as the girl, where would he sit to balance the seesaw—at A, B or C ? (11)



Here are three cog wheels, A, B, C.
Which cog wheel turns in the same direction as A. Is it B, C or neither ? (12)

Which cog wheel turns the fastest ? Is it A, B or C ? (13)

Is a flower alive ? (yes, no, don't know) (14)

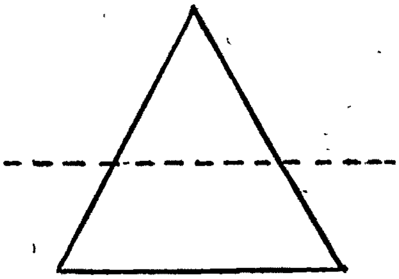
Are all things that move living ? (yes, no, don't know) (15)

A needle is stuck upright in a cork. The cork floats on water.
Draw two lines showing position of cork and needle. (16)

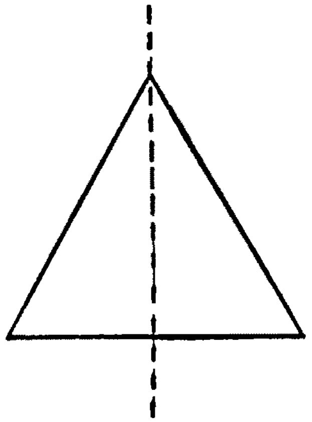
A circle cut from cardboard is held edge on to an electric light. The shadow of this cardboard falls on a screen. Draw the shadow you would see on the screen. (17)

A rubber ball is cut through the middle. Draw the shape made by the cut. (18)

The top of a cone is cut off as in the diagram.
Draw the shape made by the cut. (19)

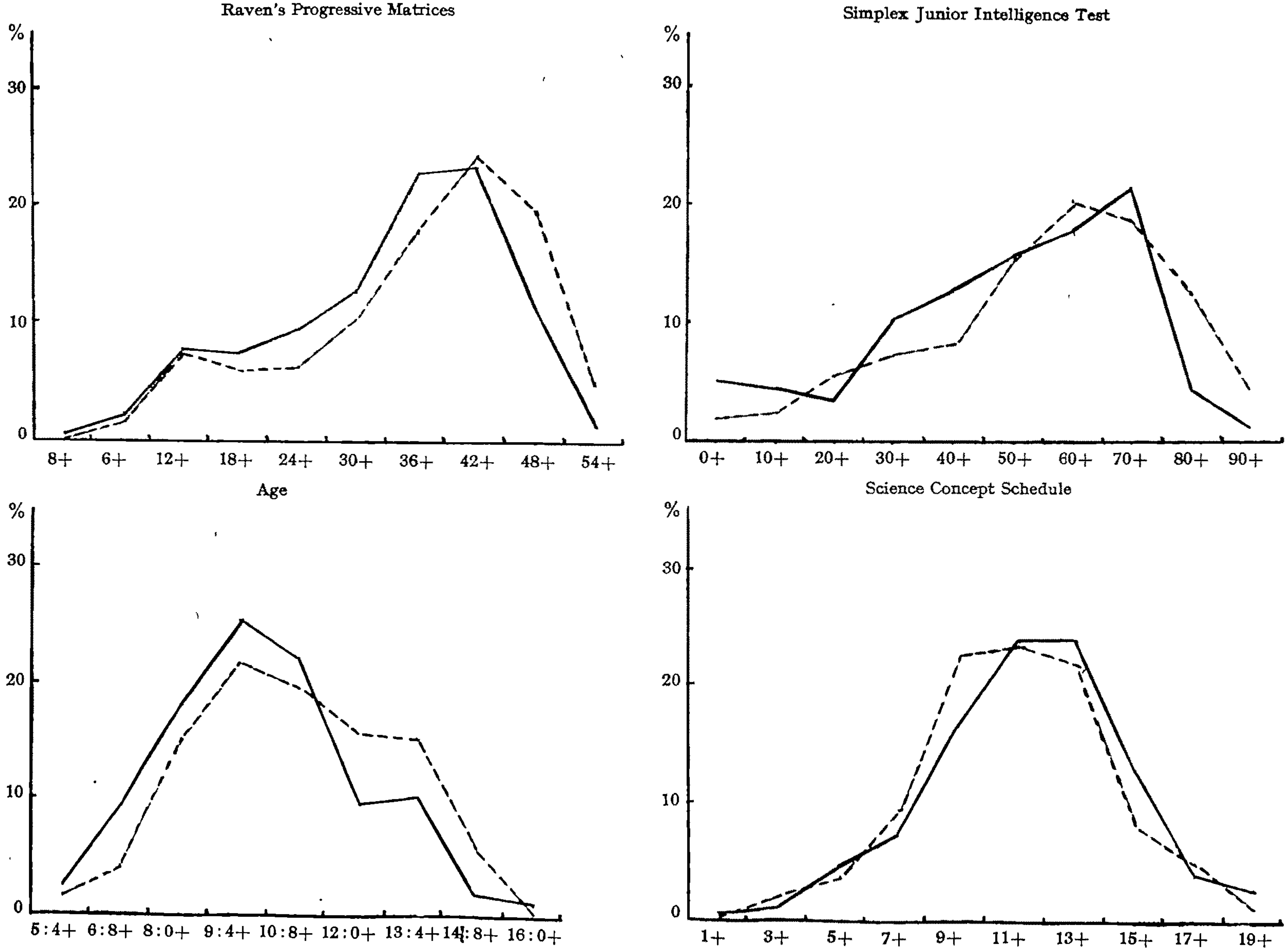


A cone is cut in half, as shown in the diagram.
Draw the shape made by the cut. (20)



IX.—APPENDIX 3.

Distribution of the boys' and girls' scores on the four variables.



Raven's Progressive Matrices

Simplex Junior Intelligence Test

Age

Science Concept Schedule

LEADERSHIP IN A DUAL CULTURAL SETTING: A SOCIOMETRIC STUDY OF CLEAVAGE BETWEEN ENGLISH AND AFRIKAANS-SPEAKING SCHOOL CHILDREN AND THE ROLE OF THE LEADERS IN BRIDGING IT

BY R. K. MUIR
(*Education Department, University of the Witwatersrand*)

SUMMARY. Moreno's sociometric test technique was used to measure cleavage in the school-leaving classes of four parallel-medium high schools. On the basis of an analysis using sociograms and statistical tests, the two culture groups in the schools were found to co-operate in the environmental organisation of the schools, although many friendships between the groups were not made. The leaders were found to initiate contact in the environmental organisation of the schools and it is suggested that this contact leads to friendship ties.

I.—INTRODUCTION.

SOUTH AFRICA is faced with several peculiar problems in the field of human relations. Not only is there the problem of Black and White races, but the White race itself is composed of two groups which have different cultures. The most marked outward manifestation of the two cultures is seen in the two languages, English and Afrikaans, spoken by the groups. In the schools the pupils of the two White groups have, by and large, been kept apart, and some educationalists, politicians, and indeed people from all ranks of society, have felt that by putting pupils from the two groups together to be taught in the same school, one step will have been taken towards unifying the White population. The assumption is that a more intimate knowledge of the members of the other group will lead to a liking for the group. Research has shown that while this is often the case, increased knowledge of another group can lead to antipathy and even stronger feelings of identify for one's own and rejection of the other (Herman and Schild 1961, Sherif and Sherif 1953). In any society in which there are sharply defined differences between groups composing the society, as in South Africa, there will have to be liking born of more immediate contact between the groups if harmony is to be achieved.

Apart from separate educational provision for the two cultures, dual medium and parallel-medium schools have been tried. The former suffers from the waste of time in teaching and learning resulting from using two languages in the same classroom and the critics of the latter say that there is no contact between the groups anyway, so they are also a waste of time. This study was designed to examine the measure of cleavage in the structure of English and Afrikaans-speaking pupils who attend parallel-medium schools. Further, the aim was to examine the rôle of the leaders in this dual cultural setting.

II.—SOCIOMETRIC TEST TECHNIQUE AND GROUP STRUCTURE.

A modified version of Moreno's sociometric test technique was used (Moreno 1953). The test made it possible for members of different cultural groups to choose each other, so cleavage between the cultures could be measured by the choice pattern between them. A choice expressed by a member of one culture group for a member of another has been called a cross-cultural choice.

The technique has limitations and some assumptions underlie its use. Sociometric status is the combined score of all the choices received by the individual on the criteria in the test. This objective measure is indicative of how the group evaluates each of its members. Those with a score of nought are called isolates and those with a statistically high score are called stars. Various interpretations may attach to star status and it can only be explained by reference to the social setting. The test technique merely identifies stars, but, generally speaking, stars are important to the group to ensure its continued existence and to uphold its cultural traditions. Stars act as spokesmen, initiate and direct action, set the values and goals to which the group aspires, and determine how they should be realised. Research has suggested that stars are better liked and tend to be highly evaluated in areas other than those which were the basis of their status. Further, they interact more with their fellows (Hurwitz, *et alia*, 1955). While it is important for the leaders to uphold the group's way of life—and if they transgress too far from their position, they will lose their status—they are, nevertheless, given more latitude in their behaviour than are the members of lower status. They are expected to mix with other groups in upholding the prestige of their own, and in doing so, they interact with the members of those groups.

Interaction between members is expressed in a sociometric test in choice patterns and cannot be defined further than the pattern. A choice made by a member *A* of another *B* means that *A* has a desire to interact with *B*. *B* has also made choices and should *B* choose *A* it would seem that this mutual choice situation means that there is interaction between *A* and *B*. There is at the least a desire to associate on the part of both, and this has probably already led to successful interaction which has reinforced the desire.

The structure of the group becomes the sum total of types of choice patterns and the distribution of choices. When structure is considered the emphasis is not so much on numbers of choices but on their configurations. Why an individual should be accepted, why the members should choose in a particular pattern, or why the group should be so structured are questions which the sociometric technique alone cannot answer. Other methods of study would need to be developed to answer these questions.

This study was of school populations in which cleavage could be expected between the English and Afrikaans-speaking pupils which composed the population, and the sociometric test has been found well suited to reveal cleavage in the structure of a group (Criswell 1947, Loomis 1943). However, it cannot explain the cleavage or say why cross-cultural choices were made in certain cases, although the pattern revealed is such that interesting conclusions can be drawn from it.

Each member of the test population was asked to make sociometric choices of friends and of leaders in the school standard, every opportunity being given for cross-cultural choices to be made. Choosing of friends can be regarded as inter-personal capacity for participating in the intensely personal organisation of the group, which Jennings called the ' psychegroup ' (Jennings 1947) and Homans along similar lines called the ' internal structure ' of the group (Homans 1951). Each has suggested a wider pattern of contact in the more formal outward organisation of the group—called by Jennings the ' sociogroup ' and by Homans the ' external structure '—and leadership is a phenomenon of this structure.

Although their approaches differ and the setting of their studies is different, the hypotheses formulated by both Jennings and Homans suggest that where two cultural groups impinge on one another, there will be more contact between them in the leadership structure of the group. Conversely, the friendship structure will be structured within the culture group with much less inter-group contact. Homans made a further hypothesis that interaction in the external structure will lead to interaction in the internal structure.

It was the aim of this study to see how these hypotheses held true for groups whose members belonged to different cultures. These two hypotheses were formulated using Jennings' terminology:

Hypothesis 1. When two groups with different cultures are brought together in the formal organisation of school life, leaders will be chosen irrespective of the group to which they belong, but friends will be chosen on an in-group basis. Cleavage will be marked in the psychegroup structure but not in the sociogroup structure.

Hypothesis 2. The leaders in the groups of English and Afrikaans-speaking pupils will interact in the sociogroup structure, and this interaction will result in interaction in the psychegroup structure. Thus, the leaders will bridge the cleavage between the groups.

III.—THE RESEARCH METHOD.

After a preliminary trial, these four criteria were chosen for a final questionnaire:

(1) With which of the pupils in your standard would you like to remain in contact in the years ahead?

(2) Which of the pupils in your standard would you most like to invite to your home to enjoy leisure activities?

(3) Many schools given an award to the pupil who has served the school best. Qualities of leadership and ability to put school interests before personal interests are important considerations for the award. Which of the pupils in your standard do you think worthy of the award?

(4) Which of the pupils in your standard do you think have shouldered or could shoulder the responsibilities of prefectship best?

The first two criteria, referred to as the friendship criteria, were of a general nature. By 'general' is meant that a general area for activity is indicated in the criterion rather than a specific field of interaction; the person is chosen for his personal qualities and the kind of person he is rather than for any specific skill of which he is capable. It was desired to elicit choices which would reflect the pupils' desire to maintain a permanent pattern of group structure, so the first criterion was found very suitable. Both criteria were suitable for eliciting choices which would imply a desire to interact socially, however limited the social milieu of all or some of the members of the group may have been. The criteria were meaningful to all and did not leave any but the anti-social in an embarrassed position of being unable to make the choices. There was every opportunity in these criteria for reciprocal choice and mutual association.

The last two criteria, referred to as the leadership criteria, were specific in nature. The basis for interaction was limited, in that in the choice of leaders, the choice process tends to favour a few individuals, and ignores a large number of them. The leadership criteria had meaning for all the experimental groups because the choice of pupils on the basis of these criteria is made annually by the school staff in order to present awards and prizes. The choice arouses great interest in the pupils and being given an opportunity to express a viewpoint was readily accepted.

The two types of choice criteria reflect activity in the psychegroup and sociogroup structures. The friendship criteria require choices on the basis of personal satisfaction. The choice is made of those who have personal attraction for the person choosing, and there is the opportunity for the choice to be reciprocated. The process is two-way in that the choosers are expressing choices of each other, and so the reciprocal choices are revealed. The leadership criteria demand choices not on the basis of personal qualities of attraction for others, but rather on the basis of the contribution the leader can make to the group life, on his importance to the group and his ability to get things done. This choice process is one-way. All choices tend to go to the few, and the few tend to choose each other, so there is not the same opportunity for mutual choice. It is only the leaders who interact in a sociometric sense.

The number of choices allowed on each criterion was five. School-leaving pupils have sufficiently developed social contacts to make this number of choices easily and studies have shown that five choices provide the most stable sociometric results (Gronlund 1959). A fixed number of choices also has pratical advantages in statistical interpretation. Only positive choices were required, as it was felt that resentment would be caused by asking for rejections. The spaces were given on the questionnaire for the choices to be filled in them.

A carefully-prepared Afrikaans version of the questionnaire, in which it was ensured that the meaning and intention of the original English version was retained, was available for the Afrikaans-speaking members of the experimental groups.

The questionnaire was administered with both a verbal and a written set of instructions in both languages, care being taken to establish rapport and to bring the members of the two cultures together so that cross-cultural choosing should be possible, without making it seem desirable. The four experimental groups with a total of 171 subjects were distributed in sub-groups which were designated as follows:

Sub-group *a*—English-speaking boys.
Sub-group *b*—English-speaking girls.
Sub-group *c*—Afrikaans-speaking boys.
Sub-group *d*—Afrikaans-speaking girls.

Sub-groups *a* and *c* were very small in numbers in three of the experimental groups and for the purposes of this article these sub-groups have largely been omitted from the discussion.

The first experimental group consisted of the matriculation class of a high school, whereas the other three experimental groups attended commercial high schools in the Transvaal. The school-leaving classes were chosen as the members of the classes had known each other for a minimum of two years and some pupils had been together for most of their school life. This ensured that a stable and clearly-defined group structure had developed. Further, the leaders in the school community are found among the seniors and a hierarchy of leadership status is most evident among them.

The schools were not randomly selected and the sub-groups were not well balanced in numbers, but as unilingual schools are the order of the day, this was unavoidable.

In the experimental groups, three variables for division were operating: cultural differences, sex differences and the division of the groups into classes for instruction. To care for the last variable a control measure was used. In one experimental group a cultural group was divided into two classes for instruction. The cleavage recorded in this case could be compared with that recorded when other variables were operating.

IV.—THE ANALYSIS OF THE RESULTS.

Sociomatrices for each experimental group were prepared, the choices on friendship and leadership criteria being recorded separately. Both the results of the sociometric testing and the social situation in each school were examined for evidence of two group structures. There was strong evidence of them, particularly in the social setting, but inter-correlations of the choices received on leadership and friendship criteria ranged from ·73 to ·77. The status of the stars was next examined to ensure that they were leaders in their own in-group, otherwise any conclusions based on their cross-cultural choosing would have been spurious.

Sub-groups *a* and *c*, the English and Afrikaans-speaking boys, were included in the analysis, but as six of these sub-groups consisted of between two and five individuals, outgroup choices were forced upon them and a pattern of choosing was revealed different from that obtained with the more typical girls' groups. Friendship choices were directed to boys rather than girls, but these special circumstances and sex cleavage, which is known to exist in the choice of friends, must be borne in mind. For these reasons the results for the girls' groups only have been considered in the following discussions.

Two types of analysis were undertaken: (*a*) the reciprocated and partially reciprocated choices were graphically depicted by means of sociograms and (*b*) statistical tests were applied to determine the level of significance of the volume of cross-cultural choices recorded.

(*a*) *The Sociograms.*

Each sociogram consisted of four concentric circles with the members identified by numbers and arranged within the circles according to their sociometric status and so as to show their interactions with their fellows most clearly. Bronfenbrenner's (1945) fixed frame of reference was used, and apart from one case where the numbers of choices which were made fell so low that statistical limits also needed to be lowered, those members with sixteen choices and more appeared in the centre, those with more than ten in the next concentric circle, those with four and more in the third and those with less than four in the outermost ring. These limits range from the ·02 to the ·05 level of significance, depending on the numbers in the experimental groups. The girls were identified by a number placed in a small circle and the boys by a number placed in a triangle, and then the reciprocated choices (mutual pairs) were shown by a solid line and partially reciprocated choices by a dotted line. The English-speaking members of the test population were placed in the upper segment of the circle and the Afrikaans-speaking in the lower. All cross-cultural reciprocations were, therefore, easily seen by the lines crossing these segments.

The sociograms drawn on the basis of friendship criteria gave a clear diagrammatic picture of psychegroup structure (see Fig. 1 for an example. Three isolate boys are not shown). Those in the centre were highly chosen and the complex network of solid and dotted lines connecting them with each other and with those less chosen show that they made the most satisfactory interactions. The stars had as many as six of their ten choices reciprocated either on both criteria or on one, and to have five choices reciprocated wholly or partially was common. Those placed nearer the second circle were connected with the stars and members of like sociometric status in triangular patterns. However, the number of reciprocations declined and those in and near the outermost circle had fewer reciprocations still. Those who were significantly underchosen obviously yearned for interaction with the more important members of the group and not with each other. The lines expressing mutuality of choice sent out arms from the centre embracing fewer and fewer members as the outer sections of the sociogram were reached. In the outermost circle there was a chain pattern without the complexities of the triangle and clique patterns so prevalent in the centre. There was no successful cross-cultural interaction as expressed by reciprocal choices for Experimental Group I, although eighteen cross-cultural choices were recorded. In the other experimental groups there were ten occasions when cross-cultural interaction was recorded, and these interactions involved fourteen members. Of these, six were stars on friendship criteria, none was significantly underchosen, and on only two occasions was a star not a party to the reciprocation. They maintained their status on leadership criteria.

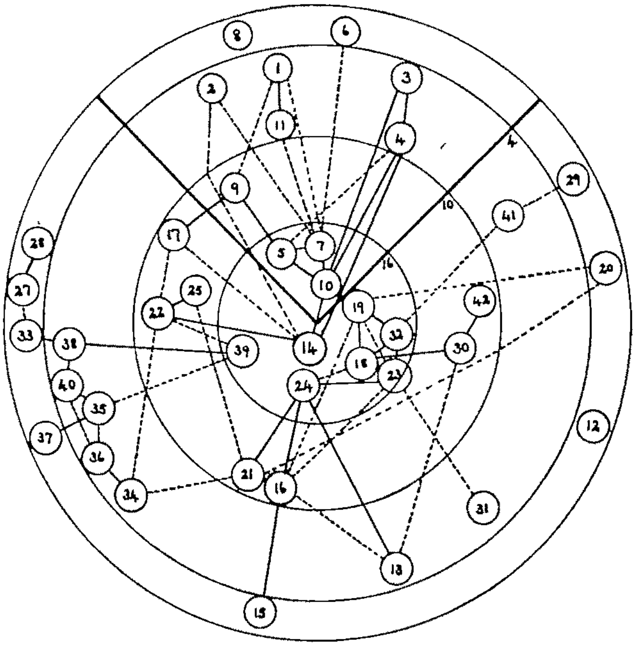


FIGURE 1

AN EXAMPLE OF A SOCIOGRAM DRAWN ON FRIENDSHIP CRITERIA

Upper segment—English-speaking group.
Lower segment—Afrikaans-speaking group.
Solid lines indicate complete reciprocation.
Dotted lines indicate partial reciprocation.

On leadership criteria the difference between members of high and low sociometric status was very marked (see Fig. 2 for an example). The stars interacted, but those outside the innermost circle were desultorily included, and large numbers were in complete isolation in the outermost circle. This is explained by the nature of the criteria and the tendency for the choice situation to be one-way. However, in choosing themselves the stars did not recognise

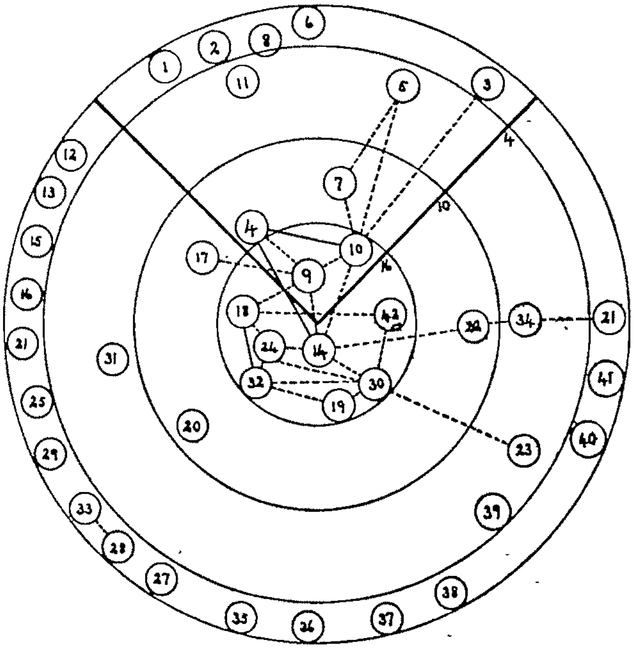


FIGURE 2

AN EXAMPLE OF A SOCIOGRAM DRAWN ON LEADERSHIP CRITERIA.

Upper segment—English-speaking group.
Lower segment—Afrikaans-speaking group.
Solid lines indicate complete reciprocation.
Dotted lines indicate partial reciprocation.

cultural barriers. In all four experimental groups, there was cross-cultural interaction among the stars and a rejection of the remainder of the population.

Of the total of twenty individuals involved in making cross-cultural choices, sixteen were stars and the remaining four had a choice status little short of that of a star. A comparison was made of individuals who had achieved star status and interacted cross-culturally on leadership criteria with those in a similar position on friendship criteria. Among the twenty in this category on leadership criteria were included all six who were in this category on friendship criteria and eleven who achieved star status on friendship criteria.

The control group showed a cleavage between classes but also between two cliques on friendship criteria, and the cliques were not based on class division. The leaders of each clique were in one class and drew followers from both. This clique pattern and class cleavage disappeared with leadership criteria, which indicates that the leaders bridge cleavage whatever the variable producing cleavage may be.

From the evidence of the sociograms these conclusions would seem to emerge:

(1) Cross-cultural interaction is greater in the sociogroup structure than in the psychegroup structure of the groups.

(2) Cross-cultural interaction is initiated by the leaders. This is particularly so in the sociogroup structure.

(3) The fact that cross-cultural interaction in the psychegroup and sociogroup structures was largely confined to the leaders would suggest that contact in the latter group structure leads to contact in the former structure. This would be a reasonable conclusion, although conclusive proof of it cannot be given.

(*b*) *A Statistical Analysis.*

The volume of choice directed from one cultural group to another indicates a willingness to interact, and as the sociograms depended on mutual choices which were not encouraged by the nature of the leadership criteria, other methods of interpreting the results were sought. Criswell's double-ratio method (1943) was used to measure in-group preference and sub-group *b's* relation to *d* and *vice versa* produced consistently lower indices for leadership criteria. This index indicates the drawing power of the in-group and the bigger the index the greater the degree of self-preference. Therefore, the self-preference indices confirmed that in the sociogroup structure there was not as much preference for the in-group as there was in the psychegroup.

However, a statistical design was wanted which would test the chance occurrences of differences in the choosing pattern, and overcome the difficulties commonly met with in sociometric results of skewness of the distribution and inflation of the significance if number of choices given or received is made the basis for analysis. A test which ignored the fact that each individual was permitted several choices on each criterion and which grouped choices together as if each choice came from a different individual would give spuriously inflated levels of significance. The individual making the choices had to be taken as the basic unit.

Question 1.

The first question which it was desired to answer was whether there was more out-group choosing on leadership criteria for each cultural group. The hypothesis was that cross-cultural interaction was greater in the sociogroup structure than in the psychegroup structure of the groups.

For each of the psychegroup and sociogroup structures an index of the individual's in-group and out-group choosing was found. This preference index was found for each of the two group structures using the formula

$$1\mathrm{P} = \frac{\mathrm{I} - \mathrm{O}}{\mathrm{O} + \mathrm{I}}$$

where I and O represented the choices, the individual made to the in-group and the out-group, respectively. Two scores for each individual were thus obtained, one score IP_F derived from the choices made on friendship criteria and the other score IP_L derived from choices made on leadership criteria. It will be seen that both IP_F and IP_L scores can range from $+1$ (the in-group wholly preferred) to -1 (the out-group wholly preferred). The difference, *d*, was then obtained for each individual, thus:

$$d = \mathrm{IP}_\mathrm{F} - \mathrm{IP}_\mathrm{L}$$

The members of each experimental group were then placed in one of three categories according to their *d*-values being greater than, equal to, or less than zero. The results are shown in Table 1. It will be seen that for all groups there are more individuals in category 1 with *d* greater than zero than in Category 3 with *d* less than zero. This would suggest that the preference for the in-group is greater for the friendship than for the leadership criteria—which is the hypothesis under examination. A test of the significance of the differences between the two sets of frequencies is provided by the Binomial Expansion, from which one can derive the probabilities of obtaining frequency differences as great as those between Category 1 and Category 3 in Table 1. The probabilities appear in the last column of the Table.

TABLE 1

DISTRIBUTION OF MEMBERS OF GROUPS ACCORDING TO THE DIFFERENCE, *d* (BETWEEN THE PREFERENCE INDICES ON THE FRIENDSHIP AND LEADERSHIP CRITERIA) BEING GREATER THAN, EQUAL TO OR LESS THAN ZERO.

Experimental Group	Sub-group	Category 1 $d>0$	Category 2 $d=0$	Category 3 $d<0$	Significance of difference of frequencies in Cat. 1 and Cat 3 expressed as a P-value.
I	b	3	6	0	·24
	d	9	9	1	·02
II	b	11	3	3	·06
	d	8	4	0	·02
III	b	11	0	0	·002
	d	19	10	1	·0002
IV	b	10	1	0	·002
	d	12	6	1	·004

TABLE 2

DISTRIBUTION OF STARS AND NON-STARS ACCORDING TO NUMBER OF OUT-GROUP CHOICES MADE.

Experimental Group	Out-group choices made			χ^2	Level of Significance
		0—1	2+		
I	Stars	2	4	2·42	N.S.*
	Non-stars	17	15	—	—
II	Stars	2	6	·0004	N.S.
	Non-stars	7	14	—	—
III	Stars	2	7	·2	N.S.
	Non-stars	12	20	—	—
IV	Stars	2	5	2·24	N.S.
	Non-stars	16	7	—	—

* Not significant.

For experimental groups III and IV high levels of significance were obtained while two further sub-groups, one from each of the other two experimental groups, provided significant P-values. It can, therefore, be safely concluded that there is more cross-cultural interaction in the sociogroup than in the psyche-group structure of the experimental groups.

Question 2.

The second question which was asked was whether the stars gave out-group choices more readily than the non-stars on the leadership criteria. Leadership choosing only was examined for this purpose and the distribution of out-group choosing for the stars was tested against that for the non-stars. A 2×2 contingency table was formed for each experimental group by grouping the stars into two categories—those who had made one or no choices from the out-group in the one category and those who had made two or more choices among the out-group in the other. The non-stars were arranged in the same categories and the significance of the difference between the two distributions was tested by a Chi-squared test. In calculating the values of chi-squared for the different 2×2 tables, Yates' correction was applied when any of the expected values fell below five. It was not necessary to apply the exact method for 2×2 tables since the chi-squared values in all cases were either highly significant or not at all so. The results are shown in Table 2 where it can be seen that the non-stars made out-group choices as readily as the stars.

Question 3.

The third question which was asked was whether the stars received out-group choices more readily than the non-stars. By definition, the stars receive significantly more choices than non-stars, but in this test out-group choices only were examined to compare the acceptability to the in-group of out-group stars and out-group non-stars. A 2×2 contingency table was formed. The same procedure was followed as for question 2, but the choices which stars and non-stars received formed the basis of the distributions. In Table 3 it can be seen that the distributions thus formed had a 1 per cent. level of significance for three of the experimental groups. This means that the stars not only achieved their status in the in-group, but also were recognised as stars by the out-group.

TABLE 3

DISTRIBUTION OF STARS AND NON-STARS ACCORDING TO NUMBER OF OUT-GROUP CHOICES RECEIVED.

Experimental Group	Out-group choices received			χ^2	Level of Significance
		0–1	2+		
I	Stars	2	4	12·69	1%
	Non-stars	22	0	—	—
II	Stars	1	7	13·05	1%
	Non-stars	19	2	—	—
III	Stars	1	8	16·32	1%
	Non-stars	28	4	—	—
IV	Stars	4	3	·14	N.S.
	Non-stars	17	6	—	—

The answer to questions 2 and 3 is that out-group stars attract choices from both stars and non-stars in the in-group. Stated another way, there is no difference between stars and non-stars in the giving of choices to out-group members, but these choices are given only to the stars in the out-group.

V.—CONCLUSIONS.

Both the sociograms and the statistical tests showed that the first hypothesis can be accepted with one reservation. Leaders were chosen cross-culturally, but the leader of the in-group was preferred. There was not a complete absence of cleavage in the socio-group structure. That the leaders bridge the gap between the culture groups as stated in the second hypothesis can be accepted for the sociogroup structure only. Further, the evidence of the sociograms suggests that there is the strong possibility that interaction in the sociogroup structure leads to interaction in the psychegroup structure, although no adequate proof can be given of this.

VI.—THE RELEVANCE OF THE RESULTS IN THE SCHOOLS.

These results indicate that given the opportunity of associating in parallel-medium schools the pupils do co-operate in the environmental organisation of the school. They need to play games together, not as members of opposite teams; they need to work together as prefects, or as members of school clubs and societies ; and they need to take part in the same plays, speech competitions parades and so on. Particularly, the leaders require to be brought into contact as they bridge the cleavage. In South Africa there are several parallel-medium schools with a proud tradition and those who know them are convinced that differences can be a source of strength rather than weakness, as out of diversity comes growth. Furthermore, after school days, the two cultural groups work together, frequently inter-marry and show goodwill in many ways.

On the other hand, the single medium school has become the declared policy of the political party ruling at present and separation is being perpetuated even in military training after school. Without the lead given by the top administrators in education, and without the extra effort needed in organisation, administration, and even in tact, to make parallel-medium schools a success, they are doomed to failure. As in the school populations, in the national life it is the leaders who must initiate contact between the groups, and without it separation and isolation will be perpetuated.

ACKNOWLEDGEMENT.—The writer wishes to acknowledge guidance in statistical interpretation received from Professor N. Gourlay, Head of the Department of Education, University of the Witwatersrand.

VII.—REFERENCES.

[1] BORGATTA, E. F , and BALES, R. F (1953). Interaction of individuals in reconstituted groups. *Sociometry*, 16, 302-320

[2] BRONFENBRENNER, U. (1945). The measurement of sociometric status, structure and development. *Soc. Mono., No* 6.

[3] CRISWELL, J. H. (1943). Sociometric methods of measuring group preferences. *Sociometry*, 6.

[4] CRISWELL, J. H. (1947). A sociometric study of group integration. *Sociometry*, 10, 259-267

[5] GRONLUND, N. E (1959). *Sociometry in the Classroom*. New York : Harper and Brothers

[6] HERMAN, S. N., and SCHILD, E. O. (1961). The stranger group in a cross-cultural situation. *Sociometry*, 24.,

[7] HOMANS, G. C. (1951). *The Human Group*. London : Routledge and Kegan Paul, Ltd.

[8] HURWITZ., J. I., ZANDER, A. F , and HYMOWITCH, B. (1955). Some effects of power on the relations among group members. In Cartwright, D., and Zander, A. F. (Ed.), *Group Dynamics · Research and Theory*. London : Tavistock.

[9] JENNINGS, H. H. (1947). Leadership and sociometric choice. *Sociometry*, 10, 32-49.

[10] JENNINGS, H. H. (1947). The psychegroup and sociogroup. *Sociometry*, 10

[11] JENNINGS, H. H. (1950). *Leadership and Isolation* (2nd ed.). New York . Longmans Green.

[12] LOOMIS, C. P. (1943). Ethnic cleavage in the South-West as reflected in two high schools. *Sociometry*, 6.

[13] MORENO, J. L. (1953). *Who Shall Survive ?* New York : Beacon House.

[14] NORTHWAY, M. L. (1953). *A Primer of Sociometry*. University of Toronto Press.

[15] PROCTOR, C. H., and LOOMIS, C. P. (1951). Analysis of sociometric data. In Jahoda, M., Deutch, M., and Cook, S. W. (ed.), *Research Methods in Social Relations, Part* 2 : *Selected Techniques*. New York : The Dryden Press.

[16] SHERIF, M., and SHERIF, C. W. (1953). *Groups in Harmony and Tension—An Integration of Studies on Intergroup Relations*. New York : Harper and Brothers.

(*Manuscript received December 8th*, 1962.)

THE LEARNING EFFECTS OF PICTURES*

BY OLOF MAGNE AND LENNART PARKNÄS
(*University of Gothenburg, Sweden*)

SUMMARY. The present study concerns the problem of the learning effects of projected pictures when used as instructional aids and the validity of measurement techniques in media research.

Two hypothetical functions of pictures are discussed: The ' information value ' and the ' motivation value ' of the instructional material.

In order to study the information value of pictures, three experiments were designed to test the following hypotheses :

The method of testing defines which part of the retention (verbal or pictorial) we intend to measure, and

A method of testing, adequate to the mode of presenting the information, covers a larger part of the total retention than a method that is inadequate to this mode.

Different materials were constructed in order to give the subjects verbal or pictorial information and to measure retention. The results supported the hypotheses. At the same time the experiments suggested that, under the circumstances studied by the authors, pictorial learning was superior to verbal learning.

I.—INTRODUCTION.

A LARGE number of authors have demonstrated that learning effects are increased if visual aids are used.

A survey of the literature dealing with this field of research will be found, for example, in the works of Hoban, Dale and Finn (1949), Hoban and van Ormer (1951) and May and Lumsdaine (1958). Moreover, it may be mentioned that the effect of instruction material is discussed in teachers' guides, such as those by Dale (1954), Wittich and Schuller (1957), the two year books concerning audio-visual methods (1949, 1954), published by the National Society for the Study of Education, and a very large number of works on teaching individual subjects such as geography, biology, mathematics, and physics.

The authors of the present investigation have interested themselves in the effectiveness of projected pictures when used as instructional aids. There have been comparatively few studies on this subject. In general, the findings show that pictures are passably effective, provided the teacher instructs the pupils how to study the content of the pictures (Buswell 1935). Other investigations, however, have led to exactly the opposite conclusion. That pictures produce only insignificant learning effects, at least in some cases, is shown, for example, by Vanderer (1950), Stampolis (1953), Duregard (1953), Vernon (1951, 1953, 1954), and Klingberg (1957). Jerkedal (1955) has pointed out the importance of testing the children's knowledge by an adequate method in relation to the information given.

In Vernon's first experiment (1951), two texts, each comprising about 400 words, were used ; one of these was not illustrated, and the other had graphic diagrams. Vernon suggests that the graphic illustrations obstructed the retention of the text and that this result indicates that the subjects (girls aged 16–18) lacked the ability to profit by the knowledge the diagrams were intended to convey.

---

* Translation into English by W. H. Hilton-Brown, Stockholm.

In Vernon's second experiment (1953) four short articles on two different topics were used; and in each case the material was presented in two versions. One version was written in a popular style and was illustrated by pictures that were expected to arouse interest. The other version was more concise and factual, and contained graphic illustrations and tables. The subjects (boys and girls, aged 16–18) were required to recall the contents orally and to answer certain questions. Vernon was unable to observe any appreciable difference in retention between the two articles.

A third study by Vernon (1954) was carried out with subjects who were 12 years of age. Also in this case, the illustrations appeared to exert only an insignificant influence.

Klingberg (1957) has conducted an investigation in accordance with Vernon's studies. Klingberg interprets the results as follows: (1) retention of texts with pictures is about the same as retention of unillustrated texts; and (2) pictures hinder children from reading a text with their full attention.

II.—MEASUREMENT OF LEARNING EFFECTS OF ILLUSTRATIONS.

These conflicting results raise the question whether it is clearly understood *what* is being measured. What should be the criteria for the choice of illustrations? The question can be extended to cover all instructional material. If we disregard the social aspect of education and concentrate on the acquisition of facts, we must judge entirely by the effectiveness of the instruction as defined by the degree of retention a certain time after learning. The degree of clearness has been employed as a rule of thumb for the appraisal of effectiveness. This outlook has been systematized in what is known as the cone of experience, a diagram which we reproduce from Dale (1954) (Fig. 1).

Some authors have maintained that the cone form indicates a reduction in learning effect with increasing abstraction.

Although this approach is based on teaching experience, we are convinced that learning effects are not dependent merely on the degree of clearness. We may imagine a verbal representation such as a newspaper article and a first-hand experience in the form of an interview, both having the same content, where the first yields greater retention than the second. Consequently, there seem to be other functions (or 'dimensions') of the instructional material which influence effectiveness. We consider that the degree of information about the object is one of the functions of the instructional material; and we refer to this function as the '*information value.*'

If we return to the example of the newspaper article, and the interview, it can be asserted that their information value is the same. Should we find different degrees in retention, this must be due to another function or 'dimension.' Thus, the pupils who do the interviewing may, for some reason, have attained a positive or negative attitude to the material. There seems to be a qualitative (as opposed to a quantitative) or an illustrative (in contrast to an informative) aspect of the learning material. With regard to this group of factors, we speak of the '*motivation value*' of the instructional material.

With this conception as a starting point, in experiments on the learning effects of illustrated and unillustrated texts, we ought to avoid using pictures with a high motivation but a low information value and then drawing general conclusions concerning learning with the aid of pictures in contrast to learning without pictures.

Furthermore, with regard to the problem of measurement it hardly appears defensible to measure the retention of pictures only by verbal tests.

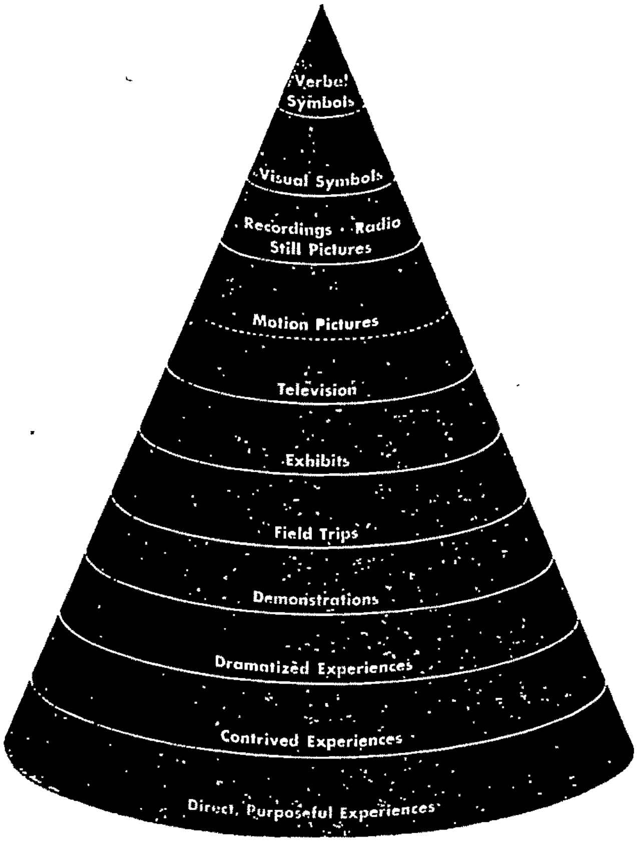


FIGURE 1
THE " CONE OF EXPERIENCE."

We presuppose there is yet another factor that has to be taken into account, viz., the aim of learning. If we wish the pupils to give an adequate account of retained visual concepts, visual types of learning material should be selected; but, if we wish an account in the form of a first-hand experience, the mode of learning should be a first-hand experience; and finally, if the pupils have to acquire verbal knowledge, the material to be studied should preferably be presented in a verbal form.

III.—HYPOTHESES.

Thus, in order to study more closely this complex of problems, we have designed three experiments to test the following hypotheses:

(1) The method of testing defines which part of the retention we intend to measure.

(2) A method of testing, adequate to the mode of presenting the information, covers a larger part of the total retention than a method that is inadequate to this mode.

If these hypotheses are confirmed, a comparison may be made between verbal and pictorial learning or between different combinations of these two modes. Comparisons of this kind will give a more detailed conception of the relative effectiveness of the various learning methods.

Experiment 1.

In Experiment 1, the authors aimed at making comparisons between learning with words and with pictures. Though these comparisons were to be as direct as possible, they were, however, to resemble actual teaching situations.

The learning material (a set of colour slides) was projected on to a screen in the classroom; the subjects had to study each slide for 3 minutes. Each slide represented four plants, and of every combination of four plants, there were two versions: one consisting of a coloured picture of the four plants (see Fig. 2) and the other of a verbal description. The test consisted in judging whether, in a second slide the petals, leaves or roots of the plants had been changed. There were two versions also of the test picture: . a 'pictorial' (see Fig. 3) and a verbal representation.

Thus, information could be furnished verbally or pictorially, and retention could be measured either by verbal or by pictorial means, i.e., there were four experimental conditions scheduled in Table 1.

TABLE 1

EXPERIMENTAL CONDITIONS.

Type of Information	Type of Test	Designation
Picture	Picture	PP
Words	Words	WW
Picture	Words	PW
Words	Picture	WP

The subjects were 228 boys and girls aged 13. Each class was divided into two halves so that the average marks for science should be the same.

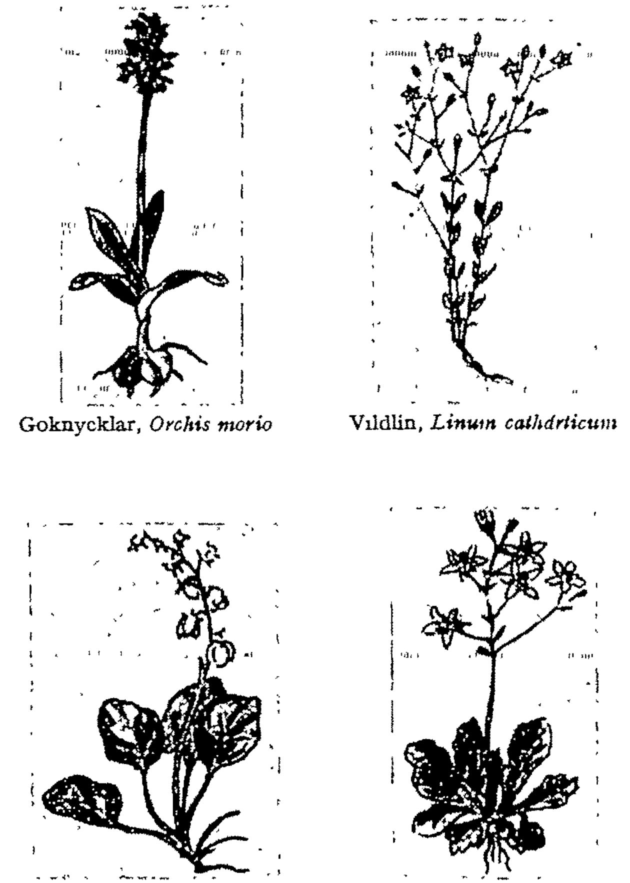
Each group was shown a slide with information, and immediately afterwards, the corresponding test slide. The slides for the verbal versions were designated W_1, W_2, W_3, and W_4: and for the pictorial versions P_1, P_2, P_3, and P_4. One half of a class was given the following information series: P_1, W_2, P_3, W_4, and the other half was given: W_1, P_2, W_3, P_4. This applied to all the classes. The test slides were varied in different combinations. For each experimental condition (PP, WW, PW, WP) there were four tasks, which will be referred to as items. Item 1 was composed, for instance of P_1W_1, P_1P_1, W_1P_1, and W_1W_1.

Retention was determined by assigning plus scores for correct answers.

The results were subjected to a statistical analysis comprising variance analysis, t-testing and χ^2-testing of each item. Table 2 gives a summary of the p-values for the different significance tests.

Our results suggest that there are differences between the various experimental situations which cannot be due to chance.

Moreover, in three out of our four items we can show that pictorial learning is superior to verbal learning, when retention is measured by both pictorial tests and verbal tests. In item III, no significant difference was observed between the modes of presentation. If retention is measured only by pictorial tests,

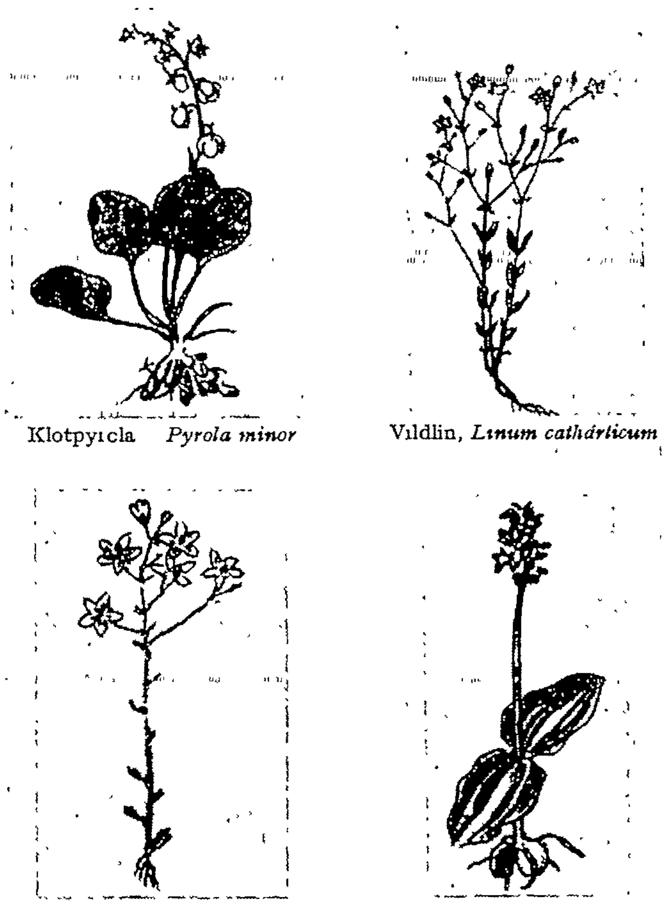


Goknycklar, *Orchis morio* Vildlin, *Linum cathárticum*

Stjarnbräka, *Saxifraga stelláris* Goknycklar, *Orchis morio*

FIGURE 2

Information Slide P_1.



Klotpyrola *Pyrola minor* Vildlin, *Linum cathárticum*

Stárnbracka, *Saxifraga selláris* Goknycklar, *Orchis morio*

FIGURE 3

Test Slide P_1.

(The slides in the experiment were in colour.)

TABLE 2

P—VALUES FOR THE DIFFERENT SIGNIFICANCE TESTS.

(The experimental situation which afforded the highest mean value is mentioned first, unless otherwise stated).

Number of Item	I	II	III	IV
Variance Analyses :				
Between experimental situations	†††	†††	†††	†††
Mode of presentation (P × W)	†††	†	0	†††
Type of test (P × W)	†††	†	0	0
Interaction	†††	†††	†††	†††
t-testings :				
PP—PW	†††	†††	0	0
PP—WP	†††	†††	†††	†††
PP—WW	†††	†††	0	†††
WP—PW	0	0	0*	0*
WW—PW	0	†	†††	0
WW—WP	0	†	0	0*
χ^2-testings :				
(PP+PW) × (WP+WW)	†††	0	0	†††
(PP+WP) × (PW+WW)	††	0	0	0
(PP+WW) × (PW+WP)	†††	†††	0	†††

* PW had the highest mean value
††† Indicate significance on the 0·001 level.
†† Indicate significance on the 0·01 level.
† Indicate significance on the 0·05 level.

pictorial learning is found to be superior to verbal learning. In this experiment we obtained differences for all our four items ; these differences indicated the superiority of pictures. There is only a 0·001 probability of this result being due to chance. If, on the other hand, retention is measured only by verbal tests, we found that verbal learning was superior although this was not very pronounced.

Furthermore, we can report tendencies and, in two cases, significant differences which indicate that the picture tests cover a greater part of the total retention.

Moreover, for all items, it was found that a higher degree of retention was recorded when tests adequate to the mode of presentation (i.e., PP or WW) were used than when the tests differed from the information type (i.e., PW or WP).

In order to ascertain what the strongly significant interaction, which occurred in all items, may indicate, we shall study item II in a four-fold table, where the frequencies are the sums of the plus scores.

TABLE 3

SCORES OF ITEM II.

	Information		Difference
	P	W	
Test P	206	144	62
Test W	142	170	—28
Difference	64	—26	

From Table 3, it is apparent that the difference between the retention of pictorial information, measured by pictorial tests and by verbal tests (64) is greater than the difference between the retention of verbal information measured by pictorial and verbal tests respectively (–26). Had there been no interaction, both these differences would have been about the same. Similar results were obtained with regard to the difference between the retention of pictorial information measured by pictorial tests and verbal information measured by pictorial tests (62) ; and the difference between the retention of pictorial information measured by verbal tests and verbal information measured by verbal tests (–28). The direction which the differences tend to take is also apparent. It is interesting to note that PP (frequency 206) is superior to WP (frequency 144) to approximately the same extent as it is superior to PW (frequency 142) ; and that WW (frequency 170) is similarly superior to WP to about the same extent as it is to PW. From a practical point of view, this means that we do not obtain any difference between the retention of information given pictorially and measured verbally and the retention of the same information given verbally and measured pictorially. This was apparent also from the t-testing, where we found that PW tended to predominate in two cases, and WP in two other cases ; but there were no significant differences.

The argument indicated in the preceding section applies also to I and III, and, tc a great extent, to IV.

In short, our results may be said to support the hypothesis that the method of testing influences the definition of retention. If it is intended to compare pictorial verbal knowledge or to test the retention of pictorial and verbal information, then retention should be measured by both pictorial and verbal tests.

At the same time, the experiment verified that, in the case studied by us, pictorial learning was superior to verbal learning.

Experiment 2.

The learning material in Experiment 1 was used in an experiment with a somewhat different design, where the subjects took part in a diploma course in education at the University of Göteborg.

The design of this experiment was as follows : The participants were divided into four groups. In Experiment 2, each group formed two teams by lot, in such a way that both instruction (pictorial and verbal instruction, respectively) and tests (pictorial tests or verbal tests) were varied in the groups as well as in the teams. The time of exposure for the instruction pictures was 2 minutes, whereas for the test pictures there was no time limit.

Table 4 summarizes the comparisons between the four experimental conditions in Experiment 2.

The sequence of the experimental conditions, in accordance with the size of the mean, was the same in Experiments 1 and 2, with the exception of verbal information tested by pictorial tests (WP) and pictorial information tested by verbal tests (PW) ; but in both cases the difference between these tests was insignificant.

Furthermore, we found that in both experiments, the mean retention of the pictorial information was greater than that of the verbal information.

What is of the greatest interest, however, is the more important result that, even among adults, the means for PP and WW are higher than those for PW and WP. If the retention of the pictorial and the verbal information had been

TABLE 4

MEANS AND t-TESTING OF DIFFERENCES BETWEEN MEANS IN EXPERIMENT 2.

Experimental Conditions	Designation	No. of subjects	Mean	t for differences		
				WW	WP	PP
Pictorial information—pictorial tests	PP	20	14·50±0 45	2·39†	4·24††	4·60††
Verbal information—verbal tests	WW	20	12·95±0·47	—	1·58	2·16†
Verbal information—pictorial tests	WP	21	12·00±0·38	1·58	—	0·76
Pictorial information—verbal tests	PW	20	11·55±0·45	2·16†	0·76	—

†—Significance at 0·05 level.
††—Significance at 0·001 level.

tested only by the pictorial test, a significant difference would have been obtained which could have implied that the retention of the pictorial information was greater than that of the verbal information. A similar investigation, where only the verbal test was applied, would have shown that verbal information was more effective than pictorial information. Thus, quite different conclusions can be drawn, depending on the test method employed.

Let us now imagine a practical learning situation where the information is presented both verbally and pictorially. The following two cases will then be of interest:

(1) The retention of this combination of pictorial and verbal information is tested by pictorial tests. On the basis of Experiments 1 and 2, we may expect retention to be at least as high as that of pictorial information alone.

(2) The retention of the combination of pictorial and verbal information is tested by verbal tests. In this case, retention should be greater than when pictorial information is tested by verbal tests, and, in accordance with the results of Vernon's and Klingberg's investigations, the degree of retention should be about the same as that of verbal information alone. The combined retention measured by verbal and pictorial tests should then be found to be higher than the retention of verbal information.

In order to test these hypotheses a third experiment was conducted.

Experiment 3.

The information consisted of a lesson on Greenland given both with and without a filmstrip. The teachers who took part in the experiment made a joint plan for the lesson on the basis of the script for the filmstrip and the test. The only difference between the two types of lessons was the showing of the filmstrip. Each lesson took 20 minutes; testing occurred immediately afterwards; no time limit was imposed.

In all, 192 boys and girls, aged 11-12, took part in the experiment. As instructors, three teachers were employed who had thorough experience of the age-group in question. In order to eliminate, as far as possible, any teacher effects, the teachers replaced each other in the different groups.

The test form contained five verbal questions (Nos. 1–5) and five pictorial questions (Nos. 6–10).

The results of the testing with the verbal items 1–5 are summarized in Table 5.

TABLE 5

SCORES ON VERBAL ITEMS.

Experimental group	Score		Total	
	0	1		
Classes without a filmstrip	98	362	460	$\chi^2=0{\cdot}11$
Classes with a filmstrip	101	349	450	$0{\cdot}70<P<0{\cdot}80$

Thus, no statistically significant difference was observed between the groups taught with a filmstrip and those taught without one, when *retention was measured by a verbal* test. Our results agree well with Vernon's and Klingberg's observations, and also with the hypothesis formulated by us.

The corresponding results of testing with pictorial items 6–10 are shown in Table 6.

TABLE 6

SCORES ON PICTORIAL ITEMS.

Experimental Group	Score		Total	
	0	1		
Classes without a filmstrip	135	325	460	$\chi^2=34{\cdot}74$
Classes with at filmstrip	59	391	450	$P<0{\cdot}001$

There was a significant difference in retention between the group taught with a filmstrip and that taught without one when *retention was measured by a pictorial test*. This is in agreement also with the hypothesis formulated by us ; on the other hand, it does not agree with the findings of previous investigations.

Hence, in this experiment, also we were able to demonstrate that the estimated retention depends to some extent on the character of the tests. Consequently, it is necessary to be careful when drawing conclusions with regard to learning with and without pictures, if retention is measured only by verbal tests or only by pictorial tests. Naturally, the validity of our results is restricted to the experimental situation studied by us. We would point out, however, that this resembles very closely real classroom situations. It is, of course, impossible to draw any direct conclusions concerning the knowledge acquired by the reading of texts with and without illustrations, but there is much to indicate that this problem is similar to the problems studied by us.

We would summarize this experiment as follows: When applying verbal tests, there were only small differences between the two groups of children. On the other hand, it appeared that retention, tested by pictorial items, was greater for the groups taught with a filmstrip than for those taught without one. Consequently, we are apparently justified in concluding that, in the situation we investigated, teaching with filmstrips yielded a higher degree of combined retention than did teaching without filmstrips

IV.—DISCUSSION.

It has been said that the best and most effective form of teaching is to give the pupils direct first-hand experience. Naturally, this is not always possible. At times, this can be done in Physics, Chemistry, Biology, and Mathematics. In certain circumstances, it may also be arranged in vocational schools and in connection with practical work in workshops But, if knowledge is to be acquired of foreign languages or foreign countries and cultures, this has to be communicated by the aid of pictures and recorded speech. It is in order to communicate this type of knowledge that audio-visual aids are required. Though they represent an indirect form of knowledge, they are regarded as more effective than verbal symbols.

In what is known as the cone of experience, which was previously referred to, reality is represented by the base, and the farther we proceed from this, the farther we are removed from reality. The cone represents an approach, where a direct and simple causal relation has been assumed to exist between proximity to reality and teaching results. Naturally, such a simplified representation of the effects of teaching methods lends itself readily to criticism,

Instead of the cone of experience, we suggest a model where the information and motivation values of the medium are the decisive factors. If we suppose these values to be located on the axes of a co-ordinate system, then a number of experimental situations become self-evident. A series of such experiments would confirm the validity of this model. Then we would find, for specific teaching units, the combination of motivation and information values that yields maximum retention Let us consider the following cases respecting the amount of retention:

(1) Varying the motivation value while keeping the information value constant (teachers' ratings may be used for grading the motivation value).

(2) Varying the information value, motivation value constant.

(3) By simultaneously varying the information value and the motivation value, it would be possible to determine the conditions under which a maximum of retention is obtained.

We leave it an open question whether the negative parts of the axes of the hypothetical model correspond to actual conditions and whether the axes of the model are to be assumed to form right or oblique angles, i.e., if information and motivation are correlated or not.

If the acquisition of knowledge is regarded from this standpoint, it will be admitted that there are direct, 'pictorial' teaching-methods which are ineffective and verbal methods which are effective in respect of learning.

There is, however, a further objection to the approach that has given rise to the cone of experience. We have been able to show that it is not defensible to measure the retention of various types of learning with tests representing only one type of material. The investigation indicates that pupils should be allowed to study pictures, if pictorial retention is required. On the other hand, if primarily verbal retention is desired, verbal material should be employed; and furthermore, if first-hand experience is the objective, the pupils should be given first-hand experience.

V —REFERENCES.

BUSWELL, G. T. (1935). *How People Look at Pictures.*

DALE, E (1954). *Audio-visual Methods in Teaching.*

DUREGARD, A. (1953 : 7). Geografiundervisningens metodik. *Folkskolan*, 68-76, 99-108.

HOBAN, C. F., DALE, E., and FINN, J. D. (1949). Research on audio-visual materials. *The 48th Yearbook of National Society for the Study of Education, Part I.*

HOBAN, C. F., and VAN ORMER, E. B. (1951). *Instructional Film Research*, 1918-1950.

HOVLAND, G. I., LUMSDAINE, A. A., and SHEFFIELD, F. D. (1949). *Experiments on Mass Communication.*

JERKEDAL, A. (1955 : 9). Tillägnandet av kartkunskaper i folkskolan. *Folkskoklan*, 43-49.

KLINGBERG, G. (1957 : 11). Värdet av bilder vid inlärning av läroboksstoff. *Folkskolan*, No. 7, 7-10.

MAGNE, O E. F., and PARKNAS, L. (1960). Inlärningseffekten av bilder. *Skola och, samhalle*, 2, 55-71.

MAY, M. A., and LUMSDAINE, A. A. (1958). *Learning from Films.*

SCHULLER, C. F. (1954). *The School Administrator and his Audio-visual Program.*

STAMPOLIS, A. (1953). Effectiveness of filmstrips in teaching economics. *Education*, 74, 115-120.

VANDER MEER, A W. (1950). Relative contributions to factual learning of pictorial and verbal elements of a filmstrip. *School Review* (February, 1950), 84-89.

VERNON, MAGDALEN D (1951). Learning and understanding. *Quarterly J. Exp. Psychol.*, 3, 19-23.

VERNON, MAGDALEN D. (1953). The value of pictorial illustrations. *Brit. J. Educ. Psychol.*, 23, 180-187.

VERNON, MAGDALEN D. (1954). The instruction of children by pictorial illustration. *Brit. J. Educ. Psychol.*, 24, 171-179.

WITTICH, W. A., and SCHULLER, C. F. (1957). *Audio-visual Materials, Their Nature and Use.*

(*Manuscript received* 13*th December*, 1962.)

PERSONALITY FACTORS AND SCHOOL ACHIEVEMENT

A COMPARISON OF BRITISH AND AMERICAN CHILDREN

BY

H. J. BUTCHER

(*Department of Education, University of Manchester*)

M. AINSWORTH

(*Sharston County Secondary School*)

J. E. NESBITT

(*Department of Education University of Manchester*)

SUMMARY. Samples of 12-14 year-old British and American children were tested with Cattell's HSPQ, and their mean scores compared on fourteen personality factors and two second-order factors (extraversion and anxiety). Significant differences were found on a number of factors and, in particular, the British children were found to be less sociable, more assertive and less conscientious. The pattern of prediction of school achievement was compared, and the same factors were in general found to be related to school achievement in both countries.

I.—INTRODUCTION

ALTHOUGH it is widely recognised that personality factors play a part in determining school attainment, methods of assessing them have in general proved disappointing.

Middleton and Guthrie (1959), for example, say that " attempts to improve prediction by using non-intellective factors such as interests and personality traits have yielded quite discouraging results . . . the principal difficulty is probably heterogeneity of the criterion and non-summative or non-linear predictions."

It must be admitted that neither questionnaires nor objective tests of personality have yet reached the stage of acceptance as practical instruments of selection and guidance, while ratings are very often so bedevilled with halo effect as to be even less satisfactory. From the theoretical point of view, however, there are enough suggestive and interesting findings in the literature to make it clear that research is getting warm, so to speak, and that real effects are there to be discovered when the research instruments have been further standardised and refined. The High School Personality Questionnaire appears to be one of the more promising instruments available for use with children, and has the following advantages :

(*a*) It provides measures of fourteen primary factors, and also of two main second-order factors into which the primary factors fall, i.e. anxiety and extraversion.

(*b*) The factors measured have emerged from a considerable body of prolonged research. They are briefly described in the tables that follow. A full description is provided in the handbook of the test (Cattell, Beloff and Coan, 1958).

(*c*) The factors correspond in most cases to those that have been extensively studied in adults.

(*d*) The test is designed to be used equally in the British and American cultures, and the items contain no glaring Americanisms or Anglicisms.

The present study is concerned with the following questions. Are there systematic differences between British and American children at age 12-14 on some or all of the fourteen primary and two main second-order factors which the HSPQ is designed to measure ? What degree of prediction of school attainment is gained by the use of this test of personality ? Do the same personality factors make for school success in England as in America ?

II.—PREVIOUS RESEARCH

(i) *Cross-cultural studies of personality.*

Studies comparing British and American children in terms of personality factors appear to be scarce. Sarnoff, *et al.* (1958), compared anxiety among American and English children (American from Connecticut, English from Hendon). They were primarily interested in validating a measure of test anxiety, and their hypothesis that English primary school children would show higher test anxiety than their opposite numbers in the U.S.A. was confirmed. On a test of general anxiety, however, no significant national difference was found.

Cattell and Warburton (1961) compared the scores of American University students, British graduate students (student teachers) and British students at a College of Advanced Technology on the 16 P.F. test. On ten out of fifteen primary personality factors (excluding intelligence), differences were found between the American and British University students. In addition there were considerable differences between the two British samples (sometimes larger than the British-American differences), and these were statistically significant in the case of seven out of fifteen factors, with smaller groups being compared than in the case of the British-American comparisons. The groups were also compared in terms of the second-order factors of anxiety and extraversion and the British University students were found to be less anxious and more introverted than the American. British C.A.T. students were least anxious and most extraverted of the three groups.

(ii) *The relation of personality factors to school attainment.*

Most studies have been concerned with the two broad factors of anxiety and extraversion—introversion. Findings have been rather conflicting, and the discrepancies can probably be assigned to some or all of the following causes :

(*a*) Differences in national culture pattern, e.g. extraversion may be more of an asset in American primary schools than in British ;

(*b*) Different effects at different ages, e.g. it seems plausible, as suggested by Furneaux (1957), that introversion may be more of an asset at sixth-form and university level owing to changed conditions of work ;

(*c*) Different definitions of the variables, and in particular a blurring of the distinction between anxiety and neuroticism. (A comparison, for instance, of anxiety scores based on the HSPQ and neuroticism scores based on the Maudsley personality inventory for children would be useful in this connection.)

(*d*) Non-linear relations between the variables, e.g. Cox's (1960) finding that a middling degree of anxiety favoured school work, and Grooms and

Endler's (1960) that only with a high anxiety group did a test of personality yield any improvement in the prediction of achievement;

(*e*) Varying effects against different criteria, as in Lynn's (1957) finding that anxiety had a different effect, according as the criterion was school work in English or in arithmetic;

(*f*) Sex differences. Women have been found to have higher average scores than men on introversion in a number of studies (e.g. Anderson, 1960). It is also possible that the relation between introversion and school achievement shows a sex difference.

In view of these discrepancies, and of considerations of space, a systematic review of the literature will not be offered here. A useful survey of recent papers in this area is provided by Warburton (1962).

III.—METHOD.

(*a*) *Samples used.*

The American sample consisted of the entire seventh grade in two junior high schools in Springfield, Illinois and Paxton, Illinois. They were selected as representing contrasting types of area, Springfield being the state capital, with the school in question drawing almost exclusively on urban families, and Paxton being a small country town, to which many of the children came from the surrounding countryside. The numbers of children tested were 154 and 123, respectively. There were approximately equal numbers of boys and girls in each sample and the mean age was 12·7.

The first British sample consisted of all the first-year children in an eight-stream secondary modern school in South Manchester. The number of children tested was 230. There were 113 boys and 117 girls, and the mean age was 12·5.

The second British sample consisted of 140 children in the third year of secondary schooling. The mean age of the children at the time of testing was 14·3. The sample included equal numbers of boys and girls and was designed to be representative of the age-group in ability and of urban society in the socio-economic status of parents.

The availability of two separate samples from different schools in both the U.S.A. and in England provides useful evidence concerning intra-country as well as inter-country differences.

(*b*) *Administration of the tests.*

The three authors were responsible for administering the tests to the American samples, to the British Secondary Modern sample, and to the representative British sample respectively. At Springfield, Illinois, it was possible to administer the HSPQ to the entire seventh grade at one sitting in the school hall. In the other three samples it was given by classes. Parallel forms of the test (A and B) were administered to both American samples and to the British secondary modern sample. In the case of the other, more representative, British sample it was unfortunately only possible to administer Form A. Similarly, measures of school achievement were available for the two American samples and the secondary modern sample, but not for the second British sample.

The criteria of achievemnt available in the American schools were firstly the Stanford intermediate achievement test, Form K, which was taken by both schools, and which yields a total score, and six sub-scores, as follows: paragraph meaning, word meaning, spelling, lauguage, arithmetic reasoning and arithmetic computation. Secondly, teacher ratings were obtained on the children's behaviour in school on a graphic seven-point scale.

The criteria of achievement available in the British secondary modern school sample were the results of the annual internal school examinations in English, Mathematics, History, Geography, Science and Art. The papers were set by the heads of departments in these subjects and all eight streams of children attempted the same papers. To aim at a standardization of marking, one teacher was responsible for the marking of a particular question throughout all the streams. This procedure was not adopted specifically for this research, but was part of the annual promotion policy of the school involved.

IV.—Results

(*a*) *Comparison of the two American groups.*

Table 1 below shows for the two American schools the mean scores and standard deviation on the fourteen primary and the two second-order factors measured by the HSPQ.

TABLE 1

Mean Scores (HSPQ, Forms A and B).

	Factor	Springfield (N=154)		Paxton (N=123)		*t*	Signif. of diff.
		Mean	S D.	Mean	S.D.		
Sociability	A	11·10	3·02	11·60	2·75	1·45	n.s.
Intelligence	B	14·11	2·48	13·68	2·39	1·46	n.s.
Ego-strength	C	8·92	2·96	10·47	2·49	4·74	·0001
Excitability	D	9·60	2·98	9·35	2·62	0·73	n s.
Assertion	E	8·89	2·77	9·27	2·89	1·05	n.s.
Enthusiasm	F	9·78	2·83	10·00	2·47	0·69	n s.
Conscientiousness	G	12·56	3·18	12·15	2·91	1·12	n s.
Adventurousness	H	8·39	2·92	9·56	2·56	3·56	·001
Sensitivity	I	10·14	3·28	9·95	3·08	0·50	n.s.
Passive individualism	J	10·68	2·00	9·86	2·37	3·06	·01
Timidity	O	10·68	3·02	9·42	2·75	3·64	·001
Self-sufficiency	Q2	10·73	2·20	10·46	2·28	1·00	n.s.
Self-control	Q3	10·18	2·37	10·61	2·18	1·57	n s.
Tenseness	Q4	9·32	2·73	8·68	2·61	1·99	·05
Extraversion		69·81	10·78	73·86	10·64	3·12	·01
Anxiety		109·47	18·74	101·74	15·60	3·76	·001

The urban group (Springfield) is significantly more introverted and more anxious than the rural one. The mean difference in anxiety is particularly marked, not only in terms of the second-order factor, but also in the fact that of the five primary factors on which the groups differ, four (C–, H–, O, Q4) load anxiety, and in each case the difference indicates that the urban children score in the anxious direction.

(*b*) *A Comparison between American Children and British Secondary Modern Children.*

Table 2 shows a comparison between the total American sample and the British secondary modern school sample.

TABLE 2

MEAN SCORES (HSPQ, FORMS A AND B).

	Factor	American Children (N=277)		British Children (N=230)		*t*	Signif. of diff.
		Mean	S.D.	Mean	S.D.		
Sociability	A	11·32	2·91	9·77	2·71	6·17	·0001
Intelligence	B	13·92	2·44	11·92	2·60	8·85	·0001
Ego-strength	C	9·58	2·89	9·07	2·60	2·10	·05
Excitability	D	9·48	2·84	10·12	2·40	2·78	·01
Assertion	E	9·08	2·83	11·13	2·57	8·51	·0001
Enthusiasm	F	9·88	2·67	10·09	2·78	0·86	n.s
Conscientiousness	G	12·34	3·10	10·63	2·76	6·55	·0001
Adventurousness	H	8·90	2·84	8·53	2·49	1·56	n.s.
Sensitivity	I	10·06	3·19	10·41	2·92	1·29	n.s.
Passive individualism ..	J	10·32	2·21	10·11	2·35	1·02	n.s.
Timidity	O	10·09	2·99	11·24	2·69	4·58	·0001
Self-sufficiency	Q2	10·59	2·27	10·64	2·35	0·24	n.s.
Self-control	Q3	10·38	2·30	9·68	2·58	3·20	·01
Tenseness.............	Q4	9·04	2·69	10·24	2·46	5·26	·0001
Extraversion		71·61	11·11	68·14	11·24	3·48	·001
Anxiety		105·98	16·73	114·24	14·93	5·91	·0001

The marked difference in intelligence (B) is to be expected, since the American group covers the whole range of ability, whereas the British group is selected (or 'non-selected'), and also rather younger. In addition the British children are seen to be markedly less sociable, more assertive, less conscientious, more timid and more tense, these last two factors (O and Q4) being the principal indicators of anxiety. In terms of the two second-order factors, they are considerably more introverted and more anxious.

Admittedly, the comparison is not entirely a straightforward one, since the British sample is entirely urban and does not contain children from secondary grammar or technical schools. But on the first point, it is plain that the British sample differs markedly not only from the total American sample, but from the American urban sample, and that in every case the differences are still in the same direction.

(*c*) *A comparison between American children and a representative sample of British secondary school children.*

This comparison is shown in Table 3

Perhaps the most striking feature of Table 3 is the confirmation of the very marked and highly significant difference already found between British and American children on factors E and G. The British children appear as very clearly more assertive and less conscientious. The lower sociability of British children is also confirmed, though less strikingly. The low score of the British

TABLE 3

(MEAN SCORES (HSPQ, FORM A ONLY).

	Factor	American Children (N=277)		British Children (N=140)		*t*	Level of Signif.
		Mean	S D.	Mean	S.D		
Sociability	A	5·22	1·74	4·74	1·81	2·61	·01
Intelligence	B	7·12	1·65	6·92	1·91	1·06	n s.
Ego-strength	C	4·91	1·80	4·87	1·57	0·23	n.s.
Excitability	D	5·08	1·65	5·04	1·72	0·23	n.s
Assertion	E	4·76	1·86	5·66	1·62	5·11	·0001
Enthusiasm	F	5·44	1·66	5·64	1·67	1·16	n.s.
Conscientiousness	G	6·13	1·74	5·19	1·44	5·84	·0001
Adventurousness	H	4·69	1·83	4·61	1·94	0·41	n.s.
Sensitivity	I	4·52	1·98	4·01	1·89	2·55	·05
Passive individualism	J	5·53	1·45	5·78	1·64	1·53	n.s.
Timidity	O	5·04	1·77	5·33	1·81	1·58	n.s.
Self-sufficiency	Q2	5·17	1·61	5·49	1·68	1·85	n.s.
Self-control	Q3	4·31	1·63	4·36	1·57	0·30	n s
Tenseness	Q4	4·61	1·70	4·59	1·61	0·12	n.s.
Extraversion		36·53	11·15	35·59	9·80	0·88	n.s.
Anxiety		55·24	7·40	55·71	7·88	0·59	h s.

children on factor I (sensitivity) is perhaps rather surprising. It is also interesting that (unlike the secondary modern group) the representative sample of British children does not differ significantly from the American children either in extraversion or anxiety.

(*c*) *The relation between personality factors and school achievement.*

The product-moment correlation coefficients between the fourteen personality factors and the measures of school achievement are shown in Tables 4-6 following.

TABLE 4

THE RELATION OF SCHOOL ACHIEVEMENT TO FOURTEEN PERSONALITY FACTORS. SPRINGFIELD, ILLINOIS. N=153.

	Personality Factors													
	A	B	C	D	E	F	G	H	I	J	O	Q2	Q3	Q4
Total	17*	62†	02	–09	–22†	02	34†	–01	08	06	–04	35†	18*	–02
Paragraph meaning	15	57†	03	–08	–14	05	32†	08	–05	06	–09	35†	14	11
Word meaning	04	58†	–02	–05	–08	–01	29†	–05	–02	06	03	32†	13	–05
Spelling	25†	53†	–03	–06	–30†	16*	33†	–02	14	09	01	18*	05	06
Language	21†	53†	00	–12	–24†	02	34†	–05	13	17*	–02	22†	13	06
Arithmetic-reasoning	19*	54†	04	–11	–10	–02	31†	00	–04	01	–03	35†	12	–05
Arithmetic-computation	19*	46†	–08	01	–22†	01	36†	–04	00	03	–02	26†	03	–02
Teacher rating on behaviour record	13	18*	–10	–16*	–28†	–09	18*	13	23†	03	–16*	15	26†	06

*—Signif. at 5 per cent. level.
†—Signif. at 1 per cent. level.

TABLE 5

The Relation of School Achievement to Fourteen Personality Factors.
Paxton, Illinois. N=124.

	Personality Factors.													
	A	B	C	D	E	F	G	H	I	J	O	Q2	Q3	Q4
Total	12	51†	00	–01	–02	–08	18*	01	–20*	02	–11	15	10	–02
Paragraph meaning	19*	25†	03	–05	00	–09	17*	–02	–14	–09	–11	19*	12	02
Word meaning ..	13	47†	13	–05	06	–09	21*	05	–21*	–04	–17*	19*	09	–04
Spelling	09	45†	–05	–07	–07	–15	23†	07	–01	–03	–04	10	09	–08
Language	11	43†	–11	–03	–10	–04	14	–03	–12	–01	–09	–01	01	–02
Arithmetic-reasoning	11	41†	03	–07	02	–01	10	13	–23†	06	–17*	11	10	–09
Arithmetic-computation ..	15	32†	02	–07	00	–03	22*	08	–11	08	–21*	13	11	–12
Teacher rating on behaviour record	12	26†	–04	–18*	–14	04	16	–01	11	–12	–03	00	29†	–01

*—Signif. at 5 per cent. level
†—Signif. at 1 per cent. level.

TABLE 6

The Relation of School Achievement to Fourteen Personality Factors.
Manchester. N=230.

	Personality Factors													
	A	B	C	D	E	F	G	H	I	J	O	Q2	Q3	Q4
English	11	45†	–04	110	–01	–10	15*	–10	05	07	08	13*	–07	02
Maths.	08	34†	03	–01	–07	–01	08	–01	03	00	03	15*	–04	–06
Science	–01	33†	14*	01	–05	–03	10	–05	–14*	02	–09	14*	01	–13*
History	–01	23†	09	09	–03	–03	02	–18†	–14*	–11	–01	14*	–06	–20
Geography	–10	26†	02	00	00	–08	10	–14*	–08	00	03	16*	07	–14*
Art	–07	08	16*	–16*	04	–11	22†	04	06	12	01	06	19†	–18†

*—Signif at 5 per cent. level.
†—Signif. at 1 per cent. level.

The main features of these tables appear to be as follows :

(1) As in earlier researches, the correlations are not strikingly high. With the exception of factor B (intelligence), no factor yields a correlation of over ·4. In general they are higher for the Springfield sample than for either the Paxton or the Manchester sample. This may very possibly be due to a greater standardisation of testing conditions, since the entire Springfield group was tested at one sitting in the school hall, whereas the other two groups were tested by classes.

(2) There is some consistent pattern as between the three samples. In particular, factors G and Q2 correlate consistently with all measures of achievement in all three groups. Of the twenty-two correlations involving factor G

(conscientiousness), all are positive, and in the American groups the great majority are statistically significant. Twenty-one of the twenty-two correlations involving factor Q2 (self-sufficiency) are positive, and fourteen of these are significant. The role of these two factors is nearly equal. In the Springfield sample, both yield correlations of around ·3, in the Paxton sample around ·2, in the Manchester sample around ·1.

(3) There is remarkably little differential prediction of the various aspects of achievement. In general, the factors such as G and Q2 which clearly predict success, predict it over the whole range of criteria. Similarly, the other factors, such as A (sociability) and E– (submissiveness) which predict to a lesser degree, (also Q4 in the British sample), still do so, almost regardless of school subject. The teachers' ratings on behaviour record clearly stand in a different category in this respect, but even here intelligence, conscientiousness and self-sufficiency appear to play some part. But it is noticeable that Q3 (self-control) E– (submissiveness), and D– (lack of excitability) also correlate with behaviour rating in both samples. These correlations appear to make good psychological sense.

(4) Factor A (sociability) correlates positively (but low) with all criteria in both American samples, but yields no significant correlation in the British sample. It is possible and plausible that this represents a genuine cultural difference, though it is hard to be sure owing to the general attenuation of correlations in the British group. This attenuation may be due either to the conditions of testing or to the lesser suitability of the questionnaire to British children.

V.—DISCUSSION

One obvious question comes to mind in connection with the results reported. To what extent were the national samples typical of children in the two countries?

Some further light can be thrown on this issue by a comparison of the results obtained for the American children with norms reported in the handbook of the test. Although there is naturally some deviation of our figures from these norms, inspection suggests that the national differences found may be real ones. If one were to substitute the handbook norms in Table 2 for the figures obtained from the combined American sample, every difference from the British Secondary Modern children (on the nine factors where significant differences were found) would still be in the same direction. To take one example, a striking difference was found on factor G (conscientiousness) between the American children (M=12·34) and the British Secondary Modern children (M=10·63). The corresponding mean score of the American group on whom the test was standardised was 12·2.

Similarly for the comparison between the American children and the representative British sample on Form A only. If one substitutes the handbook norms on Form A for the mean scores obtained by the American children, the differences on the factors where significant differences were found would all be in the same direction. Strikingly enough, in the case of the two factors of assertion and conscientiousness, where very marked differences were found between the two groups, to substitute the handbook norms for the actual Illinois results would make these differences larger still. Admittedly the average difference in age between the American children and those in the second British sample (amounting to some eighteen months) impairs the comparison, but the results appear sufficient clear-cut to indicate some genuine differences.

Although it has not been an aim of this paper directly to contrast secondary modern with secondary grammar children, it is interesting to compare our findings with those of Hallworth (1961). Hallworth points out that the British Psychological Society's report on Secondary School Selection states fairly categorically that there is considerably more worry and anxiety among grammar than among modern school pupils, and questions this statement in the light of his own findings, suggesting that the opposite may be true. Our own results, in which a modern school group were significantly more anxious than American children, but in which no such difference was found for a representative sample, provide some indirect support for Hallworth's findings.

Turning to the correlations with measures of school achievement, one finds results that are promising and suggestive rather than conclusive. The coefficients of correlation are generally low, though frequently statistically significant. To a large extent this may be inherent in the method of the personality questionnaire which has long been an object of suspicion and of lively satirical attack by laymen (Whyte, 1956). A particular deficiency of the HSPQ may be the obverse of its merit, the fact that by aiming to measure a large number of factors in a comparatively short testing time, it necessarily sacrifices some precision on each factor. Thus factor B (intelligence) yields lower correlations than would be expected from a full-scale intelligence test. The HSPQ was not constructed with the sole aim of predicting school achievement, but with the much wider one of providing a general picture of the respondent's personality pattern. It might well be a useful task for future research to concentrate on those factors (such as G and Q_2) which appear to be most directly related to attainment, and to develop fuller and longer scales for the specific purpose of predicting attainment.

VI.—CONCLUSIONS

(1) In Illinois, U.S.A., urban children were found to be more introverted and markedly more anxious than rural children.

(2) Children at a secondary modern school in Lancashire were more introverted and very clearly more anxious than the Illinois children.

(3) When a second British sample, which covered the whole ability range, was compared with the Illinois group, no significant differences were found on the factors of extraversion or anxiety.

(4) The American children were found to be very significantly more sociable, less assertive and more conscientious than either British group.

(5) Apart from intelligence, two of the remaining thirteen factors measured by the HSPQ correlate fairly consistently with school attainment. These are G (conscientiousness) and Q2 (self-sufficiency). Among the American children, factor A (sociability) was the next most consistent predictor.

ACKNOWLEDGMENT.—The American results were obtained as part of a research supported through the Co-operative Research Programme of the Office of Education, U.S. Department of Health, Education and Welfare (Project number 701 (8383)), and under the direction of Professor R. B. Cattell of the University of Illinois. The writers' thanks are also due to the Headmasters of the schools in Paxton and Springfield, Illinois, and of those in Manchester for their helpful co-operation.

VII.—References

Anderson, A. W. (1960). Personality traits of Western Australian university entrants. *Austral. J. Psychol*, 12, 4-9.

Cattell, R. B., Beloff, H, and Coan, R. W (1958). *Handbook for the I P A T. High School Personality Questionnaire* Champaign, Ill. · Instit for Personality and Ability Testing.

Cattell, R. B., and Warburton, F. W. (1961) A cross-cultural comparison of patterns of extraversion and anxiety. *Brit. J Psychol.*, 52, 3-16.

Cox, F. N (1960). Correlates of general and test anxiety in children. *Austral. J. Psychol*, 12, 169-177.

Furneaux, W. (1957). *The Selection of University Students.* London . Report to the Imperial College of Science.

Grooms, R R, and Endler, N. S. (1960). The effect of anxiety on academic achievement, *J. Educ. Psychol.*, 51, 229-304.

Hallworth, H. J. (1961). Anxiety in secondary school children. *Brit. J. Educ. Psychol.*, 31, 281-291.

Lynn, R. (1957). Temperamental characteristics related to disparity of attainment in reading and arithmetic. *Brit. J. Educ. Psychol.*, 27, 62-68.

Middleton, G., and Guthrie, G. M. (1959). Personality syndromes and academic achievement. *J. Educ. Psychol.*, 60, 66-69

Sarnoff, I., Lighthall, F., Waite, R, Davidson, K, and Sarason, S A. (1958). A cross-cultural study of anxiety among American and English children. *J. Educ. Pschol*, 49, 129-36.

Warburton, F. W (1962). The measurement of personality—III. *Educ. Research*, 4, 193-206

Whyte, W. H. (1956) *The Organisation Man.* New York : Simon and Schuster.

(*Manuscript received, 4th January,* 1963.)

THE FAMILY RELATIONS INDICATOR

A PROJECTIVE TECHNIQUE FOR INVESTIGATING INTRA-FAMILY RELATIONSHIPS DESIGNED FOR USE WITH EMOTIONALLY DISTURBED CHILDREN

BY

JOHN G. HOWELLS
(Consultant Psychiatrist)

AND

JOHN R. LICKORISH
(Senior Clinical Psychologist)

(Department of Child and Family Psychiatry, Ipswich and East Suffolk Hospital, Ipswich)

SUMMARY. Most of the few tests designed specifically to investigate intra-family relationships are based on psycho-analytic concepts A picture projective technique is described, which is based upon the simple theory that a child in a friendly environment, will draw upon his own experience to describe the activities of figures in ambiguous pictures. The pictures have been drawn to the authors' specifications. They show six basic family scenes which are replicated three times. This replication provides an internal check on the consistency of the replies. The accuracy of this technique was investigated by comparing the replies given to the pictures, with the information gained through psychiatric and other interviews with the families concerned.

I.—INTRODUCTION

IT is generally agreed that a child's behaviour is largely determined by his relationships with his family and, in particular, by his relationships with his parents. Hence an assessment of the child's neurotic, or delinquent behaviour should include a thorough investigation of his family relationships. For young children, a picture projective method of investigation may be used, since interviews and techniques requiring written answers are usually impracticable. A careful examination of currently available techniques showed that most of them were based upon psychoanalytic concepts and were methods of investigating personality, Bellak (1950), rather than techniques for assessing interpersonal relationships.

There are, however, five tests designed to explore the relationships between the members of a family, but only one of these was a picture test of the ' association ' type, Rabin and Haworth (1960). The test devised by Travis and Johnston is little more than a simple collection of pictures, Buros (1953). The Family Pictures produced by Cummings (1952) consists of a set of faces to which the child gives his reactions. Family Drawings by Hulse (1951) is a graphic expression of the child's relationships and depends for its success upon the child's ability to draw adequate pictures and the examiner's skill in interpreting them correctly. The Family Relations Test by Anthony and Bene (1957) is a variation of the ' forced choice ' technique, in so far as the child is presented with a number of statements which he must then ascribe to members of his family. The other test specifically relating to the family is A Test of Family Attitudes, by Jackson (1953). This consists of only eight pictures which are drawn with blurred outlines, some of which are highly structured. Although the responses to the test are 'associative,' there is also a standard list of questions for each picture. The test does not appear to be specifically based on any particular psychological theory and the author emphasises the need for other sources of information to be used in conjunction with the test results.

Whilst each of these tests may be used to assess intra-family relationships, it seemed to us that there was room for another technique based upon empirical principles which would incorporate features not found in the tests already available. Accordingly, the Family Relations Indicator was evolved to assess the child's general relationship with his family and in particular, to explore the parent-child relationships. The technique is based upon gaining 'associative' responses to a series of twenty pictures portraying family scenes. The responses to the cards are not 'interpreted' and it is not necessary for the child to make up a 'story,' which so many children find difficulty in doing. Short responses of only a sentence, or two, may be quite adequate. The replication of the pictures provide a rough internal check on the consistency of the child's replies, as well as providing the cumulative effect of a little information from each of three pictures. Several of the above features are, so far as we know, either found exclusively in our technique, or else they appear in a combination not found in other tests.

II.—PLAN OF INVESTIGATION

Aim. The aim of the investigation was to determine whether the technique we had devised, would give a reasonably accurate account of the child's relationships with his family. In particular, we wished to discover if it would give an accurate picture of the relationships between parents and child. The 'association' method of responding to the pictures was adopted, since this allowed the child the widest range of verbal expression, whilst the pictures 'channelled' his replies along the lines of family relationships.

Theory. It was assumed that if a friendly relationship were established between the child and the Examiner, then the child would describe the figures in the pictures in terms of his own personal experience. It was considered possible that 'wishful thinking' and phantasy might distort the descriptions of his relationships. In an attempt to minimise this distortion, we decided to encourage the child to make statements about the characters, instead of trying to compose a story. In this way we hoped that the responses would be factual and in accordance with the child's real frame of reference. We further assumed that the child's responses would reflect the actual home situation, since if a high degree of rapport were obtained with the child, he would have little hesitation in revealing what he felt about his family.

Preliminary Investigation. A considerable number of pictures was tried out in the preliminary investigation, many of them were taken from books and magazines, but some were drawn especially for the purpose. After a considerable amount of experiment, it was decided that *all* the pictures must be specially drawn for this investigation and that only six basic family scenes could be presented. This initial investigation continued for about three years. We were eventually able to secure the services of a professional artist, who drew the final set of pictures to our specifications, which were ;

(1) The pictures were to be black and white pencil drawings.
(2) There was to be as little emotional expression as possible portrayed in posture, gesture, or face.
(3) Clothes were not to be obviously indicative of any particular socio-economic class.
(4) An age range of 8 to 11 years was suggested for the children, with an appropriate age for the parental figures.
(5) Interior scenes were to have a minimum of detail, but sufficient to suggest definite rooms.
(6) Furnishings were to be as socially ambiguous as possible.
(7) All scenes must be reasonably familiar to children of school age.
(8) Scenes should be as little ' structured ' as possible and dramatic postures should be avoided.
(9) The pictures were to be drawn clearly, with fairly sharp outlines, heavy shading and blurring of outlines was to be avoided.
(10) Ambiguity was to be introduced by the arrangement of the figures in the pictures, not by shading or by blurring their outlines.

The drawings were photographed and mounted on twelve-sheet board for use during the experimental investigation. A specimen card is shown in Fig. 1.

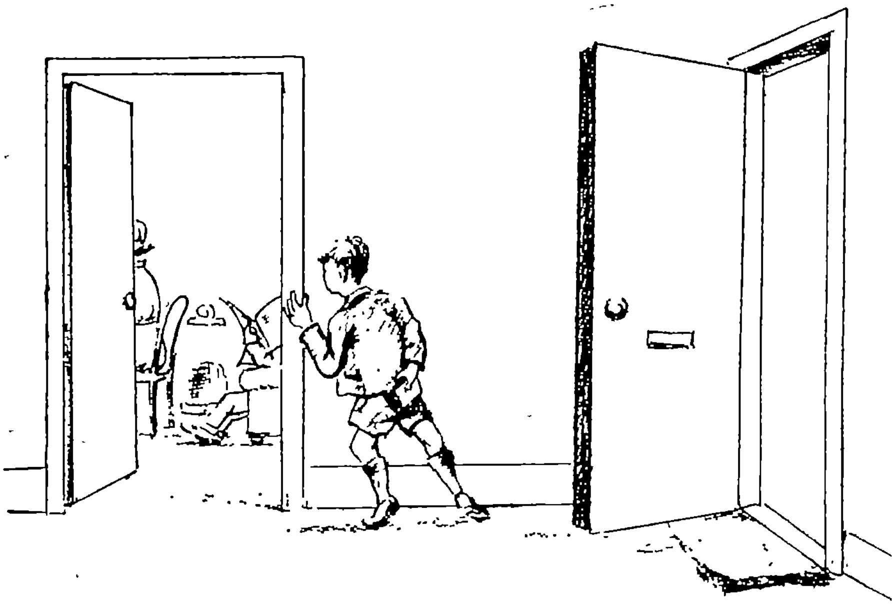


Fig. 1

Material.—The final set of pictures was made up as follows. Three introductory cards, showing family groups, were used to acquaint the child with the test procedure whilst the main series shows the following basic family situations ; parents ; father and child ; mother and child ; child alone ; a sibling group and a child and a baby. Each of these six situations is replicated three times, with the figures and the environment being sufficiently altered for the child to accept the replications as three different pictures. There is a separate series of cards for boys and girls, the only difference between them being the sex of the child in the pictures. The other figures and the environment are precisely the same in both cases. The three introductory cards are presented on each occasion, together with seventeen other cards making a total of twenty for each complete test.

Sample.—The subjects used in this investigation were children referred to the Clinic in the ordinary way, and we took fifty consecutive cases, having omitted those in which there was some unusual family constellation, or in which little was subsequently discovered about the family The ages of the subjects ranged from 4 to 17 years and there were thirty boys and twenty girls. Only thirty of the fifty children were tested for intelligence and these had a range of I.Qs. from 65 to 136. All the children and their families had been attending the Clinic for at least six months, and some of them for as long as two-and-a-half years.

Administration of Test.—The prime requirement for the test is the establishment of the maximum rapport between the child and the Examiner. This was facilitated when the test was presented on the child's second, or third visit to the Clinic. Since this investigation had to be carried out as part of the normal clinical routine, it was sometimes necessary to present the test during the child's first attendance. In these circumstances extra care was taken to ensure the establishment of good rapport. The cards were introduced to the child informally, by saying, " I have some pictures here of boys and girls and adults (or ' grown-ups ' if the child is very young). I would like you to have a look at them and see if you can tell me what you think they might be doing, or saying."

The introductory cards were then presented and very often the child would say little or nothing about them, and sometimes this reticence would extend to the initial pictures of the test series. The child was simply encouraged with praise and prompted by means of straightforward questions ; What do you think they are doing ? Where do you think they are going ? or What do you think they are talking about ? No attempt was made to cross-question the child, or to present a formal series of questions about each picture. Anything the child said was accepted, even if it were only an enumeration of the objects in the picture. When this occurred, he was praised for what he had done and then it was suggested to him that he might try to think what the people were doing.

The child's replies were recorded verbatim, sometimes in shorthand and occasionally on a tape. Afterwards they were typed in double-spacing to facilitate scoring.

III.—METHOD OF SCORING AND TABULATING THE RESPONSES

The method of scoring is not mechanical, but depends upon using psychological judgment to select from the responses those words and phrases which indicate the child's relationships with, or his attitudes towards the members of his family. In making a formal investigation of our results we scored the responses in the following manner. They were first divided into ' information units ' by means of the solidus, /. These units were sentences, clauses, phrases, or even single words. Those which described the personal relationships and attitudes of the patient were underlined, and called ' behavioural units.' Those information units which disclosed neither relationship, nor attitude were simply ignored. The ' behavioural units ' were then recorded on the Behaviour Item Sheet which was compiled as follows. From the fifty sets of responses obtained during the investigation, we selected twenty at random and analysed them. A list was made of all the expressions which were used to describe attitudes and behaviour. Some of these were found to be equivalent to one another and the redundant expressions were rejected. Forty-nine items were finally selected and grouped together as shown in the following list.

Attitudes.
1. Likes.
2. Dislikes.
3. Reluctant.
4. Tolerates.
5. Ignores.
6. Apprehensive.
7. Jealous.
8. Suspicious.
9. Sulky.
10. Sad.
11. Angry.
12. Cross.
13. Irritable.
14. Apathetic.

Verbalisations.
15. Praises.
16. Refuses.
17. Forbids.
18. Rebukes.
19. Requests.
20. Demands.
21. Defiant.
22. Threatens.
23. Blames/Accuses.
24. Argues.

Actions.
25. Initiative.
25. Plays.
27. Shows interest.
28. Helps.
29. Obeys.
30. Gifts.
31. Hits.
32. Fights.
33. Mischief
34. Injures/Accident.

Deprivations.
35 Left behind.
36. Kept in.
37. To bed.
38. No toys.
39. No food.
40. No money.
41. No playmates.

Delinquency.
42. Runs away.
43. Steals.
44. Deceitful.
45. Destructive.

Guilt Feelings.
46. Done wrong.
47. Secretive.
48. Apologises.
49. Amends

Six columns were placed parallel to these items and headed F (Father) M (Mother) B (Boy) G (Girl) Bb (Baby) and Spare.

The ' behavioural units ' were recorded by entering on the Behaviour Item Sheet the letter F.M. etc. opposite the relevant item and in the appropriate column. The letter denoted the *object towards whom* the behaviour was directed. The name at the head of the column in which the letter was placed denoted the person who produced that behaviour. If the behaviour was repeated a ' + ' was added. The column headed ' spare ' was used to record the behaviour of anyone extraneous to the family group.

If a feeling, attitude or action was expressed, without any indication of an object, then the number of the response only was entered on the Behaviour Item Sheet. When all the ' behavioural units ' had been entered in this way it was possible, by inspection, to determine the kind of relationships which existed between the various members of the family. In everyday clinical work this elaborate method of scoring is not necessary. After a little practice, the psychologist may dispense with the Behaviour Item Sheet and simply make a brief summary of the relevant items in the responses. Thus the following specimen responses may be summarised as follows :

(*a*) Father is verbally hostile to the boy. F.1.2.3.

(*b*) Father punishes boy by deprivation F.2.3. and by hitting F.2.

F.1. He is sitting up the table with Dad /and going to get something. / Father comes in, / and father ask what he was doing / and he said nothing./ *Father said, "You were taking something off the table* / and father said, *go upstairs to-day* / and *you will have no tea."*/

F.2. The little boy is standing near the fire / and father says, " What have you been doing to-day ? " / *Why are you standing near the fire ?* / *Did you play with it ?* / I didn't play with it./ " *Yes you were, you were poking it.*"/ *His father smacked him* / *and said,* " *Go up to bed* / and come down when I tell you."/

F.3. The little boy has a rocket / and father is reading his paper / and the little boy made a pile of books / and put his rocket on / *and he is telling his father what he has done.* / I have put my rocket up / and it is ready to fly. / *Father said, Put it outside,* / *don't do it in here.* / The boy took it outside / and fired it / and it was lost in the woods / *and another boy found it and took it to him.*/

In order to evaluate the accuracy of our results, it was necessary to code the information on the Behaviour Item Sheets, so that it could be handled statistically. It is notoriously difficult to code data which consist of interpersonal relationships and attitudes. Many scales have been devised, but none seemed suited to the data in this investigation. The scaling method used by Baldwin, Kalhorn and Breese (1949) was too refined for the data and the numerical coding devised by Bene (1957) was too complex. It was finally decided to employ a simple ' present-absent ' method of coding using five types of relationship which were defined as follows: (1) *Punitive action*, including any form of physical violence, or corporal punishment. (2) *Verbal hostility*, including scoldings, rebukes, " telling off " and harsh criticisms. (3) *Deprivation*, consisting of the denial, or withholding of food, pocket money, pleasures, or companionship; or being sent to bed. (4) *Indifference*. This category was used when there was an indication of a cold, non-punitive, affectionless attitude between parent and child. (5) *Positive Attitude*, indicating the presence of a helpful relationship. Because we were investigating the relationships in *disturbed* families, most of the information was entered under the ' negative ' categories. If a sample of the normal population were being investigated then more ' positive ' categories would, no doubt, be required.

IV.—EVALUATION OF THE F.R.I. RESPONSES

The information which had been recorded on the Behaviour Item Sheet was summarised, using the five categories of relationships just defined. Each of these relationships was recorded as existing, or not-existing between mother and child and between father and child. A similar assessment was made by the psychiatrist from his own knowledge of each family. The accuracy of the F.R.I. responses was then evaluated by noting the frequency with which the statements given by the F.R.I. were substantiated by the psychiatric assessment.

This is the usual method of evaluating the results of this type of investigation, as Foulds (1955) points out. The method has been severely criticised by Eysenck (1961) and Kreitman *et al.* (1961) have shown that psychiatric assessments are not always very reliable. But we had no option but to use this method, since all the cases had been investigated by one psychiatrist only and therefore a corroborative assessment was not available. The accuracy of the psychiatric assessment was probably increased by the following factors: (1) The families had been thoroughly investigated over a considerable period and the assessments were not made on the basis of simply one, or two psychiatric interviews. (2) The information gained by the psychiatrist during his own interviews was supplemented by the following sources: social workers' reports, based on interviews with the parents and visits to the home and school; results of play diagnosis; school reports; reports from the family doctor and in some cases, information gained in interviews with the probation officer, the children's officer, or the welfare worker.

There is, however, one weakness in the psychiatric assessment, which we could not avoid, since it was impossible to separate our investigation from our normal clinical practice. The children replied to the F.R.I. as part of the clinical investigation and the summary of their replies was in the psychiatrist's possession when he made his assessment of the families for the comparative study. It may, therefore, be argued that these results would contaminate the psychiatrist's assessment of traits and relationships. In reply to this we may say: (i) The F.R.I. results had been in his possession for at least six months, so that he would be unlikely to remember them individually. (ii) In making his assessment he would not, of course, refer to the psychological report. (iii) The other sources of information would be sufficiently weighty to contradict, if necessary, the results of the F.R.I.

The agreements and disagreements between the F.R.I. and the psychiatric assessment are shown in Table 1.

TABLE 1

THE FREQUENCIES WITH WHICH THE VARIOUS RELATIONSHIPS WERE SCORED 'PRESENT' AND 'ABSENT' BY THE F.R I. AND THE PSYCHIATRIC ASSESSMENT AND THE AGREEMENTS BETWEEN THEM.

Relationship	Father-Child Relationships				
	I	II	III	IV	V
					%
Punitive Action	13	17	13	7	65
Verbally hostile	28	6	8	8	78
Deprivation	14	11	20	5	74
Indifference	21	11	14	4	84
Positive Attitude	5	24	9	12	30
Totals	81	69	64	36	70

	Mother-Child Relationships				
Relationship	I	II	III	IV	V
					%
Punitive Action	19	18	7	6	76
Verbally hostile	39	2	6	3	93
Deprivation	23	5	20	2	92
Indifference	1	22	25	2	33
Positive Attitude	9	19	17	5	64
Totals	91	66	75	18	83

Col. I—No. of cases in which the F R I. and the psychiatric assessment *agree* that the relationship is *present*.
Col. II.—Agreement on the *absence* of the relationship.
Col. III.—Relationship scored 'present' by the psychiatric assessment *only*.
Col. IV.—Relationship scored 'present' by the F.R.I. *only*.
Col. V.—Ratio between Col. I and Cols. I+IV.

By using Table 1 we can estimate the extent to which the F.R.I. responses provide accurate information about family relationships. It is clear that the F.R.I. does not provide *all* the available information about parent-child relationships in the specified categories, nor should we expect it to do so. It is unlikely that the result of an investigation lasting half-an-hour (the time usually required to administer the F.R.I.) would provide as much information as an investigation extending over several months or even years. The difference between the amount of correct information provided by the F.R.I. and the total information available is indicated by Col. III. These results might be criticised on the grounds that this difference is large and that the *overall* agreements are little better than chance. But such a criticism would not invalidate the fact that the information which the F.R.I. *has actually provided* is fairly accurate as shown in Col. V of Table 1. So far as we can discover, no one has attempted to show that a picture projection method will produce *all* the theoretically available information either about individuals, or groups.

Some authors like Jackson (1950) and Walton (1959) have used the method to discriminate between nosological groups of children. Others have used this technique to provide a description of the subject's personality or attitudes, Rosenzweig (1948) Phillipson (1955) and Hartwell *et al.* (1953).

We have attempted to show that the information produced by the F.R.I. describes the relationships which actually exist within the family. Since these relationships may be independently assessed with reasonable accuracy, this assessment forms an ' objective ' standard, English (1958) against which the accuracy of the F.R.I. results may be measured. In order to do this we have compared the *total amount of information produced by the F.R.I.* with, *the extent to which this information was verified by the psychiatric assessment.* The accuracy of the F.R.I. is represented by the percentage ratios given in Col. V of Table 1.

In addition to parent-child relationships, a considerable amount of information concerning the child's *own* attitudes and feelings was given in the F.R.I. responses. The nature and accuracy of this information is shown in Table 2.

Other information of clinical value was sometimes given but not with sufficient regularity to make a further evaluation worthwhile. Occasionally references to marital difficulties, sibling, or school problems provided valuable ' clinical leads ' to be investigated during subsequent interviews.

TABLE 2

SHOWING THE EXTENT TO WHICH THE FINDINGS OF THE F.R.I. WERE SUPPORTED BY THE PSYCHIATRIC ASSESSMENT WITH RESPECT TO THE SPECIFIED TRAITS SHOWN BY THE CHILD.

Trait	I	II	III
			%
Verbally aggressive	17	13	76
Physically aggressive	27	25	92
Anti-social	24	17	71
Guilt feelings	18	11	61
Apprehensive	7	6	86
Jealous	7	4	57

Col. I —Frequency with which a given trait was present according to the F.R.I.
Col. II.—Frequency with which the F.R I result was supported by the psychiatric assessment.
Col. III.—Percentage agreement between Cols I and II.

No attempt was made to assess the reliability of the F.R.I. This is a notoriously difficult undertaking with projective techniques, since comparable sets of duplicate pictures are almost impossible to produce and any appreciable time interval between test and re-test might vitiate the reliability coefficient, since appreciable changes in the child's attitudes, or relationships, may have occurred during that interval. The 'split-half' technique might be employed in a similar manner to that adopted by Anthony and Bene (1957). But we have already incorporated a modification of this principle in our method of scoring, since a relationship was not recorded as 'present' unless it was clearly mentioned in at least two separate responses.

V.—DISCUSSION AND CONCLUSION

The Family Relations Indicator provides information about three sectors of family life. These are, the relationships between the child and father, the child and mother, and certain personal characteristics of the child itself. The types of relationships and traits which have been evaluated were not deduced from a priori principles, but were empirically derived from the results of a preliminary investigation. When the full scale investigation was carried out, no other major relationships, or traits, appeared frequently enough in the responses to warrant their addition to our original list. The preponderance of 'negative' relationships and traits in our results is due to the fact that the investigation was concerned only with emotionally disturbed children, who had been referred as patients to the Department. Naturally, not all these patients were equally disturbed, nor did they all show the same kind, or number of disturbed relationships.

It is possible that some relationships did not appear in the responses because they were too painful to describe. This would agree with the theory of the repression of painful experiences, but we were not able to investigate this point.

An inspection of Table 1 shows that the 'positive' relationship between father and child occurs in seventeen cases according to the F.R.I., but only five of these are supported by the psychiatric assessment. It is possible that this discrepancy may be due to wish fulfilment on the part of the child, who would *like* father to show him more interest and attention but, in fact, father does not do so. The existence of a gross discrepancy between the two findings on this relationship and a fairly high agreement between them on all the other relationships, tends to support this view, but there is no independent confirmation of it.

The percentage agreement for the mother-child relationships is appreciably greater than that for the father-child relationships. This may be due to the fact that most psychiatric investigations tend to be mother-child centred, with a relative neglect of father's relationship to the child. Again, there is no independent confirmation of this view.

Fantasy, which is often associated with children's projection responses, very seldom appeared in the replies we obtained. Fantasy, in the sense of an 'imaginative construction,' as opposed to wishful thinking, is easily detected when present, because the activity described is clearly impossible. Thus, a boy of 5 years 10 months replies to card M.2, " He is going to tickle Dad and then tear that book. Tear that book up and kick Dad in the tummy and get a spear and kill him."

In reply to card No. S.1, a girl aged 6 years 3 months says, " They are dancing and he kicks up his foot and hits him under the jaw and kills him." In spite of the fantasy elements in these replies the nature of the relationship disclosed seems fairly clear and we scored them as if they had been purely factual descriptions. This type of reply occurred only five times in the fifty sets of responses. Once a dream was included in a reply and in four other replies the influence of adventure stories seemed clear. Thus, a boy aged 10 years says, " The boy has a gang, and he's thinking about a plan and he dug a tunnel from one end of the playing field to the other. They made a box like that used by the men in the Horse of Troy who were escaping." However, the vast majority of replies were free from fantasy elements, and they could not have contained much wishful thinking, or the overall accuracy of the replies would not have been so great.

Responses to individual cards seldom exceeded six lines of typing on quarto paper and although this is quite short for a projection response, this defect was overcome by the replication of the pictures.

The traits listed in Table 2 were deduced directly from the replies, except that ' guilt feelings ' were inferred from the secretive, or self-punitive actions of the child.

We have found in subsequent clinical work that it is not wise to try to extract too much information from the responses. They usually show the parent-child relationships quite clearly and frequently disclose the attitude of the child towards other members of the family. In addition, the responses often provide useful information in the form of ' clinical leads,' which, although not scoreable in the manner we have described, might well be included in the psychologist's report interpreting the results. Thus, the tone of voice used in a reply, or a dramatic gesture may be very illuminating clinically, but impossible to score adequately.

We have encountered several difficulties in interpreting the responses given by the children when their family relationships were complicated by removal into foster care, the death of a parent, the presence in the home of an aunt, grandparent(s) or other relations and the presence of a much older brother, or sister, who could conceivably act as a parent figure. These complicated situations were excluded from the reported results, as there was no means of telling from the responses, whether the child was speaking of his actual parent, even if the parent had died, or whether he was speaking of a parent-substitute in the form of aunt, elder sibling, or grandparent. When a child gives a set of responses in which the characters are not clearly identified, we regard what he says about them as applicable to anyone within the family group and follow up the ' leads ' given, in subsequent interviews. We have never used a set of formal questions for eliciting information, like those given by Jackson (1952).

We endeavour to make the maximum rapport with the child and never exert pressure to obtain information. It is possible that we might sometimes gain more information by systematic questioning, but we feel that information can be obtained at too high a price and that it is unwise to press the child for information that he will not freely supply, since such pressure would lead to a loss of rapport between the child and the examiner.

It seems clear to us that a projective technique used with young children, will sometimes provide information which could not be otherwise obtained. It provides an assessment of personal and family relationships, which could be utilised for diagnostic, or therapeutic purposes. The projection method is also valuable, because it often yields information quickly, and provides valuable clinical leads at the first, or second interview. The technique we have described fulfils these functions over a limited area, but does not pretend to give extensive personality pictures, or an *exhaustive* account of inter-personal relationships.

ACKNOWLEDGMENTS.—We wish to acknowledge our indebtedness to the late Dr. W. Pawlik who assisted in the formulation of this research programme and participated in the work in its early stages. We would also like to express our thanks to Mrs. D. Garrett, Visiting Art Therapist to the Ipswich Group Hospitals for drawing the pictures to our specifications and to Mr. B. R. Reynolds for the help he gave to the project in its earlier stages.

The Family Relations Indicator has been accepted for distribution by the National Foundation for Educational Research, 79 Wimpole Street, London, W.1.

VI.—REFERENCES

ANTHONY, E. J., and BENE, E. (1957). A technique for the objective assessment of the child's family relationships. *J. Ment. Sci*, 103, 541-555.

BALDWIN, A. L., KALHORN, J., and BREESE, F. H. (1949). The appraisal of parent behaviour. *Psychol. Monogr*, 63, 4.

BELLAK, L. (1950). *Children's Apperception Test* (2nd rev. edn.). Melbourne: The Australian Council for Educational Research.

BENE, E. (1957). The objective use of a projective technique. *Brit. J. Educ. Psychol.*, 27, 89-100.

BUROS, O. K. (Ed.) (1953). *The Fourth Mental Measurements Year Book.* New Jersey: Gryphon Press.

CUMMINGS, J. D. (1952). Family pictures: A projection test for children. *Brit. J. Psychol.*, 43, 53-60.

ENGLISH, H. B, and A. C (1958). *A Comprehensive Dictionary of Psychological and Psychoanalytical Terms.* Longmans Green.

EYSENCK, H. J. (Ed.) (1961). *Handbook of Abnormal Psychology. An Experimental Approach.* New York: Basic Books.

FOULDS, G. A. (1955). The reliability of psychiatric and the validity of psychological diagnoses. *J. Ment. Sci.*, 101, 851-862.

HARTWELL, S. W., HUTT, M L, ANDREW, G., and WALTON, R. E. (1953). *The Michigan Picture Test.* Chicago: Science Research Associates.

HULSE, W. C (1951). The emotionally disturbed child draws his family. *Quart. J. Child Behav.*, 3, 152-174.

JACKSON, L. (1950). Emotional attitudes towards the family of normal neurotic and delinquent children. *Brit J Psychol.*, 41, 35-51.

JACKSON, L. (1952). *A Test of Family Attitudes.* London: Methuen.

KREITMAN, N., SAINSBURY, P., MORRISSEY, J., TOWERS, J., and SCRIVENER, J. (1961). The reliability of psychiatric assessment: An analysis. *J. Ment. Sci.*, 107, 887-908.

PHILLIPSON, H. (1955). *The Object Relations Technique.* London: Tavistock Publications.

RABIN, A. I., and HAWORTH, M. R. (1960). *Projective Techniques with Children.* New York: Grune and Stratton.

ROSENZWEIG, S., and L. (1948). *The Children's Form of the Rosenzweig Picture-Frustration Study* St. Louis.

WALTON, D. (1959). A children's apperception test—An investigation of its validity as a test of neuroticism. *J. Ment. Sci.*, 105, 359-370.

THE RELIABILITY AND VALIDITY OF THE GOODENOUGH DRAW-A-MAN TEST

A PILOT LONGITUDINAL STUDY

BY M. L. KELLMER PRINGLE AND K. T. PICKUP
(University of Birmingham, Institute of Education; Huddersfield Educational Guidance Service)*

SUMMARY. The findings reported here—which are part of a much larger longitudinal project—deal with the drawings of a man made by thirty-seven children at yearly intervals throughout their junior school career, i.e., between the ages of 7 to 10 years. A total of 148 drawings was analysed according to Goodenough's instructions. The intelligence quotients obtained from the drawings were intercorrelated: (*a*) with each other; (*b*) with the Stanford-Binet Intelligence Quotients and; (*c*) with W.I.S C. Verbal, Performance and Full Scale Quotients. An analysis of the fifty-one items of the Goodenough Scale was also carried out to investigate whether or not there is a regular, annual increase in scores over the four-year period.

The findings suggest:

(1) That for junior school children the longitudinal test-retest reliability of the Goodenough Scale is too low for it to be regarded as a stable measure of intelligence.

(2) That the Scale's validity is even lower, average correlations with quotients obtained from the Stanford-Binet Scale as well as from the W.I.S.C. Verbal, Performance and Full Scale being in the region of ·2.

(3) That its discriminating power between successive years is not entirely satisfactory.

I.—INTRODUCTION

STUDENTS of child development have for long been interested in the apparent correspondence which exists between children's maturing intellect on the one hand and the progressively greater representation of detail and proportion which is shown in their drawings of a man on the other (Cooke, 1885; Maitland, 1895; Partridge, 1902; Kerschensteiner, 1905; Lobsien, 1905; Ivanoff, 1909; Katzaroff, 1910; Rouma, 1913; Luquet, 1913; Zesbaugh, 1934). The first attempt to establish age norms was made almost 60 years ago (Shuyten, 1904). In devising a scale to assess intelligence by means of children's drawing of a man, Goodenough's (1926) approach differed from that of previous investigators in the following respects: (*a*) no *a priori* decisions were made as to features characterising intellectual ability; (*b*) artistic standards were disregarded; (*c*) efforts were made to eliminate subjective judgments as far as possible; (*d*) normative validity was derived from both chronological age and educational status; (*e*) a standard subject for drawing was chosen, but no other limitations were imposed. In all, five revisions of the scale were made, a different set of drawings being used for each revision to avoid the error of validating a point by means of the same drawings from which it had been derived. The scale now consists of fifty-one items, the norms being based upon the drawings of 2,306 children aged from four to ten years.

Since the publication of Goodenough's study numerous investigations have been carried out using the Draw-a-Man Scale. Ansbacher (1952) studying 100 fourth grade children whose mean age was ten years, found that it correlated most highly with reasoning, spatial, perceptual factors and ' tracing ' (" a test loaded with youthful willingness to follow directions.") It had least in common with verbal meaning and number. Stewart (1953) concluded " that it compares favourably in test-retest reliability with most group tests of intelligence applicable in the same age range. It also compares favourably in validity, as demonstrated on the basis of its correlations with the Stanford-Binet Test within the age groups for which it was designed, while yet possessing the advantage of being non-verbal in character." However, Stewart maintains that this instrument ought to be kept fully up-to-date; also that certain weaknesses need to be eliminated, such as unduly high norms above the ten year level and the fact that the present scoring system penalises children who draw full-face rather than profile figures.

---

* Now Director, National Bureau for Co-operation in Child Care.

Yague and Agullo (1959) studied the drawings of a man made by fifty-seven 9-year-old children. They obtained a test-retest reliability coefficient of ·846 which they considered to be too low for individual prediction. Validity coefficients between the Goodenough on the one hand and the Cattell Test I and II and Otis on the other hand, were respectively ·25, ·12 and ·38. They also concluded that the Goodenough Scale has little value in predicting school success or intelligence. Springer (1941) made an item analysis of the scale to ascertain which items differentiate between maladjusted and well-adjusted boys (as classified by teachers' ratings on the Haggerty-Olson-Wickman Behaviour Rating Schedules). Of the fifty-one items of the Draw-a-Man Test, fifteen showed critical ratios of 3·00 or higher; this was considered to be " indicative of a significantly reliable difference between the two groups that are compared." On nine of the fifteen items the maladjusted were more successful, and on the remaining six the adjusted achieved greater success. " The adjusted tend to be more successful in those items which involve motor coordination and correct proportions of arms and legs to the trunk. The maladjusted tend to include more detail in their drawings in regard to the parts of body and clothing." However, the maladjusted also showed more incongruities, especially confusions between full-face and profile presentation, and a poor sense of proportion. It is interesting to note that a somewhat similar pattern of differences was found in comparisons between normal and retarded children (McElwee 1932 and 1934, Israelite 1936, Spoerl 1940, and Goodenough 1950). Merguet (1958), having administered the scale to 1,042 Child Guidance cases ranging in age from 3 to 14 years, concluded that it is particularly useful in testing the ability of children of normal and sub-normal intelligence; whereas with children of high ability it is more meaningful as a projective technique. In Vernon's view (1960) " the published norms appear to apply pretty accurately to British children," though the reliability of re-tests is rather unsatisfactory.

The Goodenough Draw-a-Man and the Stanford-Binet Scale.

Goodenough (1926) found that for children aged from 4 to 10 years, correlation coefficients between mental age as measured by her scale and by the Stanford-Binet Intelligence Test ranged from ·56 to ·86. Williams (1935) arranged for the drawings of 100 children aged 3–15 years to be scored independently by five examiners. Correlations between examiners ranged from ·8 to ·96 and of the relationship between the Stanford-Binet and Goodenough test results, Williams stated: " While, in general, the agreement is close enough for practical purposes, extreme deviations, often of clinical significance, may be expected."

Using both Forms L and M of the Stanford-Binet Test, McHugh (1945) studied their relationship with the Goodenough Scale with a group of ninety children whose mean age was 6 years and 4 months with a standard deviation of 3·97 months. A significant correlation of ·45 was found between the mental ages obtained on the Stanford-Binet and Goodenough tests and the correlation between the intelligence quotients was ·41. Higher correlations might well have been obtained had only one form of the Stanford-Binet Test been used. With thirty-seven feeble-minded boys aged 9-18 years, Yepsen (1929) obtained a correlation of ·6.

Birch (1949) compared the Stanford-Binet and Goodenough mental ages of a group of sixty-eight borderline and mentally defective children between the ages of 10½ years and 16 years 3 months, obtaining a validity coefficient of ·62. He concluded that the Draw-a-Man test is a valid measure of mental ability in children within the age range studied whose Stanford-Binet I.Qs. are 70 or below.

Brill (1937) found that maladjusted adolescents tended to score lower on the Goodenough than on the Stanford-Binet Scale, whereas the reverse obtained for most well-adjusted subjects. " Thus a mentally defective boy who scores approximately two years or more higher on the Goodenough than is expected by reason of his Binet general mental level belongs most likely to the adjusted group. On the other hand, a mentally defective boy who scores approximately two or more years lower on the Goodenough than on the Binet belongs most likely to the maladjusted group." Hinrichs (1935) compared the Draw-a-Man performance of delinquent and non-delinquent subjects aged from 9 to 18 years, whose I.Qs. were 80 and above on the Terman. As most of them were above the age for which the Goodenough scale had been standardised, Hinrichs devised an upward extension but found this to be neither as reliable nor as valid as the part designed for younger children. The subjects were eighty-one boys living in an institution for delinquents who were matched with non-delinqent boys for chronological age and home background. The delinquents were found to have lower I.Qs. on the Goodenough than on the Binet Scale.

Using the two tests with a small group of children suffering from post-encephalitic behaviour disorders, Bender (1940) found that the I.Qs. obtained from the drawings were far below those resulting from the Stanford-Binet Test. This she interpreted as being due to a considerable disruption of the body image caused by the cerebral damage and suggested that a marked discrepancy between the two measures may indicate the possibility of some brain damage.

The Goodenough Draw-a-Man and the Wechsler Intelligence Scale for Children.

Though the Goodenough Scale was published more than thirty years ago, Sundberg (1961) found that it vies in popularity with the W.I.S.C. which was published only thirteen years ago. Indeed, among 185 testing services in the United States, it seems to occupy the highest position of any intelligence test in clinical use except the Stanford-Binet Scale. Hanvik (1953) gave the Goodenough and W.I.S.C. scales to twenty-five child psychiatric patients whose ages ranged from 5 to 12 years. The group obtained significantly lower I.Qs. on the Draw-a-Man test while the W.I.S.C. quotient fell below or equalled the Goodenough score in only four cases. The rank order correlation between the quotients obtained with the two scales was ·18. Hanvik concluded that emotionally disturbed children do not draw the human figure in a way which is commensurate with their ability as measured by the W.I.S.C.

The Draw-a-Man test used with different age groups and longitudinally.

Oakley (1940) observed that with the coming of adolescence drawings of a man undergo considerable change and there may even appear to be a deterioration in quality ; for example, lines are drawn indecisively and parts, which are difficult to draw, such as the hands, are sketchily treated or omitted. But although some rather obvious features are lost and there is a deterioration in the logical arrangement, this may be offset by a subtlety which was absent previously. In Oakley's view these changes come about through the child's losing confidence because he can now perceive how imperfect his drawings are.

Djukic (1953) examined 900 drawings made by boys and girls, ranging in age from 3 to 10 years, to determine whether there is any developmental sequence in the manner in which the various parts of the human figure appear with increasing age. He found that every part of the face or body seems to evolve independently of every other part ; moreover, the stage of development reached in representing one part does not necessarily correspond to the stage reached in representing other parts.

The earliest recorded longitudinal study is that of Luquet (1913), who kept some 1,700 drawings made by his daughter between the ages of 3 to 8 years. Care was taken to keep her drawings as free as possible from the influence of adults and other children. He found that when a new characteristic appeared it did not become ' fixed ' at once but over a considerable period of time would disappear and reappear in the drawing of the figure. McCurdy (1947) selected fifty-six drawings from a collection made by a child between the ages of 2½ to 7 years. The average Goodenough I.Q. was 171 with a standard deviation of 22. He also compared the drawings of a man made by 7-year-old children on two occasions which were separated by a three months interval ; there were fifty-nine children in the sample. The correlation between the two sets of drawings was ·69, and the difference between the mean intelligence quotients of the two samples was not significant. From these results McCurdy argued that variations in Goodenough scores rest upon fluctuation in performance and may not be due merely to an error of measurement : " The behaviour measured by the Binet test apparently has greater stability than that measured by the Goodenough test." He also concluded that the variability exhibited both by the individual and by the group is of about the same magnitude.

Ochs (1950) studied the drawings of a man made by 120 maladjusted children (eighty boys and forty girls) who had received treatment in a Children's Psychiatric In-Patient Unit. Their ages ranged from 5 to 13 years. The aim was to determine " whether improvement in social adjustment in a particular child was paralleled by increase in Goodenough drawing I.Qs. as might be expected on the basis of clinical and experimental findings." Retests were carried out between five and thirty-six months, the mean interval being fourteen months. It was found that Goodenough I.Qs. tended to increase with better adjustment and to decrease with continued maladjustment. There was little consistency regarding addition or omission of scoring items on re-tests as compared to initial test ; similarly, there were no consistent changes in the quality of the drawings. Ochs concluded that the correlation between improved social adjustment and drawing score is positive but only moderately high.

II.—DESIGN OF THE STUDY

So far there has been no attempt to investigate the reliability of the Goodenough Scale longitudinally or to validate it longitudinally against some accepted individual test of intelligence, such as the Stanford-Binet Scale or the

Wechsler Intelligence Scale for Children. The results reported here form part of a much larger longitudinal study concerned with various aspects of children's intellectual, emotional and educational development during the junior school years ; eventually data will be available for some 230 children.

In this pilot project the sample consists of thirty-seven children, twenty-one boys and sixteen girls, each of whom drew a man at yearly intervals throughout the four year period of their junior school career. This group was in no way selected but simply consisted of the children for whom the necessary intelligence test results were available at the time, namely early in the second term of their last year in the junior school. The school from which the sample was drawn serves a post-war local authority housing estate, thus consisting largely of working-class families with a sprinkling of professional people. In the first three years at this school, children are placed in unstreamed classes, but in the final year they are streamed according to level of attainment. Our sample was drawn fairly evenly from the four streams, consisting of ten A stream, seven B stream, eleven C stream and nine D stream pupils.

A total of 148 drawings, made when the children were 7, 8, 9 and 10 years old, respectively, was analysed. The intelligence quotients calculated for each of these (according to the instructions and norms given by Goodenough, 1926) were intercorrelated (*a*) with each other, (*b*) with the 1937 Stanford-Binet Intelligence quotients and (*c*) with the W.I.S.C. Verbal, Performance and Full Scale quotients. An item analysis of the fifty-one items of the Scale was also carried out to investigate whether or not there is a smooth, overall increase in scores over the four year period. Lastly, to compare the scores obtained by our experimental sample with those made by the large standardisation samples of the Stanford-Binet Test and the W.I.S.C., coefficients of correlation were calculated between scores obtained on the Stanford-Binet and the W.I.S.C. Verbal, Performance and Full Scale tests respectively ; Full Scale scores were also correlated with Verbal and Performance Scale scores. The product-moment method of correlation was used in all cases.

III.—RESULTS

There was a low degree of inter-correlation between the intelligence quotients derived from the drawings of a man made at yearly intervals over the four-year junior school period ; the highest correlation (·53) existed between the second and third year drawings and the lowest between those made in the first and second year (·22). This finding suggests that the longitudinal test-retest reliability of the scale is low, the average over the four year period being in the region of ·4 (Table I). Levels of significance reaching the 5 per cent. or 1 per cent. level of confidence are indicated in the Tables.

TABLE 1

INTER-CORRELATED I Q. SCORES OF THIRTY-SEVEN DRAWINGS OF A MAN MADE IN THE FIRST, SECOND, THIRD AND FOURTH YEAR OF THE JUNIOR SCHOOL.

	*DAM 1	DAM 2	DAM 3	DAM 4
DAM 1	—	·224	·413 (5%)	·352 (5%)
DAM 2	—	—	·534 (1%)	·431 (1%)
DAM 3	—	—	—	·499 (1%)
DAM 4	—	—	—	—

* DAM 1—Drawing of a Man, 1st year, Junior School.

Means and standard deviations are shown on Table 2. All the obtained means exceed 100, which, for an urban sample such as this, is in accordance with other findings. The mean and standard deviation of the Goodenough are lowest for the fourth year in the junior school. A mean I.Q. greater than 100 for the Stanford-Binet is also in line with other results reported for British children (Scottish Survey, 1949, Heasman 1955, Marshak 1956, Jones 1962). Similarly, the fact that the mean I.Qs. for all scales of the W.I.S.C. are above 100 but below the Stanford-Binet mean, is in line with recent findings in this country (Jones 1962). The standard deviations for both the W.I.S.C. and Stanford-Binet quotients are somewhat lower than those given for the standardisation samples at corresponding ages (Wechsler 1949 and Terman 1937).

TABLE 2

THE MEAN AND S D. OF I.QS. ON THE GOODENOUGH, W I.S.C AND STANFORD-BINET SCALES

	Mean	S.D.
Draw-a-Man, 1st year	110·4	16·0
Draw-a-Man, 2nd year	120·2	15·1
Draw-a-Man, 3rd year	117·6	14·5
Draw-a-Man, 4th year	108·5	12·2
W.I S.C. Verbal	105·0	13·2
W.I.S.C. Perf.	104·6	13·0
W.I.S.C. Full	105·6	12·1
Stanford-Binet	110·7	17·3

The inter-correlations between the quotients obtained from the drawings of a man and the W.I.S.C. Verbal, Performance and Full Scale were even lower, ranging from ·4 to ·01. They were significant (at the 5 per cent. level) only for the first year. There is a trend for the Goodenough quotients to correlate highest with the W.I.S.C. Full Scale quotients; also correlations between Goodenough quotients and all three W.I.S.C. quotients are highest in the first year and lowest in the fourth year of the junior school. A very similar pattern obtained regarding the correlation between Stanford-Binet and Goodenough quotients: inter-correlations were low throughout but highest in the first year (·37) and lowest in the fourth year (·12) of the junior school (Table 3). Thus the longitudinal validity of the Goodenough Scale seems to be low in relation to both the W.I.S.C. and Stanford-Binet Scales.

TABLE 3

CORRELATIONS BETWEEN GOODENOUGH, W.I.S.C. AND STANFORD-BINET (1937).

	DAM 1	DAM 2	DAM 3	DAM 4
W.I.S.C. (V)	·335 (5%)	·244	·215	·008
W.I.S.C. (P)	·367 (5%)	·175	·323	·079
W.I.S.C. (F/S)	·399 (5%)	·266	·313	·036
Stanford-Binet	·366 (5%)	·204	·248	·115

To investigate the comparability of the W.I.S.C. and Stanford-Binet intelligence quotients obtained by our sample with those of other validating studies, inter- and self-correlations were calculated. Correlation coefficients were found to range from ·59 to ·85 the lower obtaining between W.I.S.C. Performance and Stanford-Binet quotients and the higher between W.I.S.C. Full Scale and Verbal Scale quotients (Table 4). These results are in accordance with other studies dealing with the relationship between the W.I.S.C. and Stanford-Binet Scales. Thus they suggest that our sample is comparable to the larger samples which have been used for validating purposes, where correlation coefficients ranged from ·6 to ·9 (N. D. Sundberg in Buros, 1959).

TABLE 4

INTER- AND SELF-CORRELATIONS BETWEEN W I.S C. AND STANFORD-BINET.

W.I.S.C. (F/S) *vs.* S–B	r : ·773 (1%)
W.I.S.C. (V) *vs.* S–B	r : ·681 (1%)
W.I.S C. (P) *vs.* S–B	r . ·586 (1%)
W.I.S C. (F/S) *vs.* (V)	r : ·849 (1%)
W.I.S.C (F/S) *vs.* (P)	r : ·839 (1%)

An analysis of the fifty-one items of the Goodenough Scale indicated that there is not an overall tendency towards a regular annual increase. It would seem that with increasing age there is either a decrease in interest or in attention to detail. As the sample progressed through the junior school, there was a gradual falling off in the number of children showing steep increases in scores ; this trend was most marked between the third and fourth year and least marked between the first and second year. Over the years there was also an increase in the number of children showing a deterioration in scores ; this trend was again least marked between the first and second year and most marked between the third and fourth year. However, these tendencies did not affect the overall pattern, according to which the number of items on which there was an annual decrease was much smaller than those on which there were yearly increases.

IV.—DISCUSSION AND CONCLUSIONS

The findings of this pilot study suggest that for junior school children the longitudinal test-retest reliability of the Goodenough Draw-a-Man Scale is too low for it to be regarded as a stable measure of intelligence (around ·4). Its validity is even lower. For each of the four years, average correlations with intelligence quotients obtained from the Verbal, Performance and Full Scale W.I.S.C., as well as from the Stanford-Binet Scale, were within the region of ·2 to ·3. Hence, it would seem that the Goodenough Scale is both invalid and unreliable as a measure of intelligence for children aged seven to ten years.

In horizontal studies of drawings of a man made by children aged 7 to 11 years, Goodenough (1926) found average correlation coefficients of ·8 for validity and ·9 for reliability. More recent studies, using different methods of correlation, suggest validity to be of the order of ·4 and reliability to be around ·8. Though these figures are lower than Goodenough's, they are still higher than those found in the present study. There are at least three possible hypotheses to account for these rather divergent findings : first, that the samples employed, both in this and in most of the other studies, were too small to yield statistically meaningful results ; secondly, that the norms have become out of date, making a re-standardisation of the scale necessary ; or thirdly, that a longitudinal study is not statistically comparable to a horizontal one. Of course, these hypotheses are not mutually exclusive. When the data for our larger longitudinal study become available, they may go some way towards substantiating or disproving these hypotheses.

Our findings suggest that during the junior school period the Scale's discriminating power between successive years may not be entirely satisfactory. Some early observations seem to be borne out by our data namely that of Luquet (1913), that new features shown in a child's drawing of a man do not become 'fixed' but may disappear and reappear over a period of time, and that of Oakley (1940) regarding some deterioration in the quality of figure drawings with the onset of adolescence Nowadays quite a few children are approaching adolescence during the end of their junior school career. Thirdly, the loss in popularity of the human figure as a subject for children's drawings towards the end of the Junior school stage, was first commented upon by Maitland (1895). Studying 1,570 drawings of children who were asked to draw a subject of their own choice, she found that a human being was chosen by 45 per cent. of children aged between 5 to 7 years, by 40 per cent. of those between 8 to 10 years, by 12 per cent. of 11 to 13-year-olds and by only 5 per cent. of children who were between the ages of 14 to 17 years. In our study this increasing loss in popularity showed itself in the children's overt attitude to the task : during the first two years of this longitudinal study they settled quite happily to drawing a man, but during the third and, even more markedly, the fourth year they showed many unmistakable signs of dislike : there were sighs, groans and also delays before the drawing was started ; quite a few of the children insisted on hiding their 'man' from their neighbours as well as from the investigator (M.L.K.P.) whom by that time they knew very well.

When the Stanford-Binet and W.I.S.C. are used as validating yardsticks, it is clear that the Goodenough Scale does not yield comparable estimates of intelligence for children aged between seven and ten years. However, in the absence of a generally agreed definition of intelligence, one could argue that the Goodenough taps different aspects of ability from those assessed by means of the Stanford-Binet or W.I.S.C., especially as it is a non-verbal test. Ansbacher's (1952) findings lend some support to this view, as do some of the studies dealing with mentally subnormal, maladjusted, delinquent and otherwise handicapped children. Moreover, the Scale's greatest usefulness may well lie with younger age groups. What seems needed is a normative and longitudinal investigation with British children aged beteeen 3 to 7 years.

V.—References.

Ansbacher, H L. (1952) The Goodenough draw-a-man test and primary mental abilities. *J. Consult Psychol.*, 16, 176-180.

Bender, L (1940). The Goodenough test in chronic encephalitis in children *J Nerv. Ment. Dis.*, 91, 277-286.

Birch, J. W. (1949). The Goodenough drawing test and older mentally retarded children. *Am J. Ment. Def.*, 54, 218-224

Brill, M J. (1937) A study of instability using the Goodenough drawing scale. *J. Abn. Soc Psychol*, 32, 288-302.

Buros, O. K. (Ed.) (1959). *The Fifth Mental Measurements Year Book.* New Jersey : The Gryphon Press

Cooke, E. (1885). Art teaching and child nature. *London J. of Educ.*

Djukic, S. (1953). Evolution of certain parts of the human body and face in children's drawings. *Savremena Skola*, 8, 450-458.

Goodenough, F. L. (1926). *Measurement of Intelligence by Drawings*. Harrap.

Goodenough, F. L. (1950). Studies in the psychology of children's drawings, II.—1928-1949. *Psychol. Bull.*, 47, 369-433.

Hanvik, L. J. (1953). The Goodenough test as a measure of intelligence in child psychiatric patients. *J. Clin. Psychol.*, 9, 71-72.

Heasman, I. (1955) *Effect of Prematurity on Psychological and Physical Development*. Ph.D. Thesis, University of London

Hinrichs, W. E. (1935). The Goodenough drawing test in relation to delinquency and problem behaviour *Arch. Psychol*, New York, No 175, 82.

Israelite, J. A. (1936). Comparison of the difficulty of items for intellectually normal children and mental defectives in the Goodenough drawing test. *Am. J. Orth*, 6, 494-503.

Ivanoff, E. (1909). Recherches experimentales sur le dessin des ecoliers de la Suisse romande Correlation entre l'Aptitude au dessin et les autres aptitudes *Arch de Psych.*, 8, 97-156

Jones, S (1962). The Wechsler Intelligence Scale for children applied to a sample of London primary school children. *Brit. J. Educ. Psychol.*, 32, 119-132.

Katzaroff, M D. (1910). Qu'-est-c-eque les enfants dessinent? *Arch de Psych.*, 9, 125-133.

Kerschensteiner, D. G. (1905). *Die Entwickelung der Zeichnerischen Begabung*. Munich · Gerber.

Lobsien, M. (1905). Kinderzeichung und kunstkanon. *Z. für Padogogische Psychologie*, 7, 393-404

Luquet, G H. (1913). *Les Dessins d'un Enfant*, pp. 262. Librairie Felix Alcan, Paris.

Maitland, L. M. (1895) What children draw to please themselves. *Inland Educator*, 1, 77-81.

Marshak, M. D. (1956) *A Study of the Wechsler-Bellevue Intelligence Scale applied to British Schoolchildren*. Ph D Thesis, University of London.

McCarthy, D (1944). A study of the reliability of the Goodenough test of ingelligence. *J. Psychol.*, 18, 201-216.

McCurdy, H. G. (1947). Group and individual variability on the Goodenough draw-a-man test. *J. Educ. Psychol.*, 38, 428-436

McElwee, E. W. (1932). The reliability of the Goodenough intelligence test used with subnormal children, fourteen years of age *J. Appl Psychol*, 16, 217-218.

McElwee, E W (1934). Profile drawing of normal and subnormal children. *J. Appl. Psychol.*, 18, 599-603.

McHugh, G (1945). Relationship between the Goodenough drawing-a-man test and the 1937 revision of the Stanford-Binet Test. *J. Educ. Psychol.*, 36, 119-124.

Merguet, L (1958). The Goodenough test in child guidance. *Prax. Kinderpsychol. und Psychiatrie*, 7, 161-166.

Oakley, C A. (1940). Drawings of a man by adolescents. *Brit J. Psychol*, 31, 37-60.

Ochs, E. (1950). Changes in Goodenough drawings associated with changes in social adjustment. *J. Clin Psychol.*, 3, 282-284.

Partidge, L. (1902). Children's drawings of men and women *Studies in Educ.* (Ed. Barnes), 2, 163-179.

Rouma, G. (1913). *Le Langage Graphique del Enfant*. Misch et Thron, Bruxelles 2nd edn. pp. 14-17 and 80-83.

Schuyten, M. C. (1904). Deoorspronkelijke 'Ventjes' der Antwepsche Schoolkindern. *Paedologisch Jaarboek*, 5, 1-87.

Scottish Council for Research in Education (1949). *The Trend of Scottish Intelligence*. University of London Press

Spoerl, D T. (1940). Personality and drawing in retarded children. *Char and Pers*, 8, 227-239.

Springer, N. N. (1941). A study of the drawings of maladjusted and adjusted children, *J. Genet. Psychol.*, 56, 131-138

Stewart, N. (1953). *Review in the Fourth Mental Measurements Yearbook* (Ed. O. K. Buros). New Jersey · The Gryphon Press

Sundberg, N. D. (1961). The practice of psychological testing in clinical services in the United States. *B.P S. Bulletin*, No. 44, 1-9.

Terman, L. M., and Merrill, M. A (1937). *Measuring Intelligence*. London : Harrap.

Vernon, P. E. (1960). *Intelligence and Attainment Tests* University of London Press.

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children* (Manual). New York : The Psychological Corporation.

WILLIAMS, J H. (1935). Validity and reliability of the Goodenough intelligence test. *Sch. and Soc.*, 41, 653-656

YAGUE and AGULLO (1959). The drawing test of Goodenough and its scientific questions. *Rev Psicol. Appl. Gen*, Madrid, 14, 155-170.

YEPSEN, L. N. (1929) The reliability of the Goodenough drawing test with feeble-minded subjects *J. Educ Psychol*, 20, 448-451.

ZESBAUGH, H. A (1934). *Children's Drawings of the Human Figure* Univ of Chicago Press.

(*Manuscript received 28th December*, 1962.)

VI.—ACKNOWLEDGEMENT.

Analysis of the data was facilitated by grants from the Mental Health Research Fund and the Noel Buxton Trust.

A FURTHER STUDY OF PROGRESSIVE MATRICES (1947)

BY AUBREY J. YATES

(*University of Western Australia*)*

SUMMARY. Two groups of university students were allowed sixty minutes for Progressive Matrices (1947), indicating the point reached after forty minutes. One group began at item No. 1, the other group at item No. 21. Analysis of the data indicated that for about 13 per cent. of the sample, the score after forty minutes markedly underestimates the level the subject would be able to reach if given longer time. These subjects make few errors within the forty minute time limit, but attempt many fewer items than faster working students. It is suggested that records be scrutinized for such patterns and additional time allowed where necessary.

I.—INTRODUCTION.

IN a previous study (Yates, 1961), an item analysis of Progressive Matrices (1947) was carried out on the scores of two groups of university students. While the test discriminated very well within the groups, it was also shown that:

(i) Many of the first twenty items were of negligible discriminatory value; the middle twenty items discriminated best; the last eight items were attempted by very few subjects.

(ii) The increase in difficulty level apparently showed a sharp upward trend after item 35. However, since few subjects attempted the final items, their true discriminatory value could not be assessed.

(iii) Comparison of individual scores was rendered difficult by the fact that individuals obtaining the same score had often attempted a differing number of problems.

Consequently, it was suggested that the test might be improved if the first twenty items were omitted (while retaining the practice set) and the time limit increased from forty to sixty minutes. The present study investigated the validity of these suggestions.

II.—PROCEDURE.

The samples were drawn from a complete class of 142 external students in residence at the University of New England during the second vacation of the academic year. For practical sessions, they were split into two approximately equal groups on an alphabetical basis. At each of two practical sessions, the sub-groups were further divided into two sub-groups. After both sub-groups had completed the practice booklet, one sub-group was instructed to commence the test proper at item 1, the other sub-group at item 21. After 40 minutes' work, both groups indicated the last item solved by circling the item number. They were then given a further twenty minutes work on the test. In this way, two sub-groups submitted to each treatment were formed. Since the sub-groups submitted to the same treatment showed no significant differences in performance, results were combined to give two treatment groups: Group 3* consisted of seventy-three subjects who started at item 1; Group 4 consisted of sixty-nine subjects who started at item 21.

---

* The data pertaining to this study were collected at the University of New England, New South Wales.

At other practical sessions, all subjects were given the Advanced AL and AQ tests (described in the previous paper) as a check on the comparability of these groups, both with each other and with the external group used in the earlier experiment. No important differences were found.

III —HYPOTHESES.

The design outlined in the previous section was intended to provide evidence with respect to the following hypotheses:

(i) Omission of the first twenty items of the test would lead to a significant reduction in the number of items not attempted in the remainder of the test (items 21-48) in both the forty minute and sixty minute periods compared with the group which was given the complete test

(ii) If the first hypothesis was supported, then the increased number of items attempted would lead to an increase in discriminatory power of the test within a university sample, both for the forty minute and sixty minute periods.

(iii) The extension of the time limit from forty to sixty minutes would significantly improve the discriminatory power of the test.

IV.—RESULTS AND DISCUSSION.

Table 1 shows the average number of items passed, failed, and not attempted during the standard time limit of forty minutes, and the extended time limit of sixty minutes, for the two groups separately. In calculating these figures the first twenty items were discarded for Group 3, making their results comparable with those for Group 4.

TABLE 1

PERFORMANCE OF TWO GROUPS ON ITEMS 21-48 OF PROGRESSIVE MATRICES (1947) AFTER FORTY MINUTES AND SIXTY MINUTES

	First 40 Minutes No of Items			Total Time (Sixty Minutes) No. of Items		
	Passed	Failed	Not Attempted	Passed	Failed	Not Attempted
Group 3* $\bar{X}$	11·59	6·51	9·90	13·55	10·77	3·71
σ	4·15	4·08	5·48	3·99	5·00	4·30
Group 4† $\bar{X}$	13·01	8·99	6·00	14·30	12·58	1·12
σ	3·36	4·76	3·69	3·84	4·28	1·95

* Began test with item 1
† Began test with item 21.

---

* The groups are so labelled to distinguish them from Groups 1 and 2 employed in the previous experiment (Yates, 1961).

The principal results may be summarized as follows:

(i) Omission of the first twenty items produced a significant reduction in the number of items not attempted (9·90 vs. 6·00; $t=4{\cdot}92$, $p<{\cdot}01$). Increasing the time limit by twenty minutes also produced a significant reduction in number of items not attempted, both for Group 3 (9·90 vs. 3·71; $t=18{\cdot}34$, $p<{\cdot}01$) and for Group 4 (6·00 vs. 1·12; $t=14{\cdot}88$, $p<{\cdot}01$).

(ii) The increase in number of items attempted in sixty minutes was accompanied by a significant increase in number of items passed, both for Group 3 (13·55 vs. 11·59; $t=8{\cdot}37$, $p<{\cdot}01$) and for Group 4 (14·30 vs. 13·01; $t=6{\cdot}14$, $p<{\cdot}01$).

(iii) The omission of the first twenty items produced a significant increase in the number of items passed, after forty minutes (13·01 vs. 11·59; $t=2{\cdot}22$, $p<{\cdot}05$), but not after sixty minutes (14·30 vs. 13·55; $t=1{\cdot}13$, $p>{\cdot}05$).

(iv) The correlation between forty-minute and sixty-minute score was $+0{\cdot}950$ for Group 3, and $+0{\cdot}942$ for Group 4. These correlations should be viewed cautiously, however, since they indicate the relationship between a score X (items passed after forty minutes) and a score $X+Y$ (items passed after sixty minutes).

In order further to investigate the effects of the addition of twenty minutes to the time allowed, the following analysis was carried out. Score increases in the final period ranged from zero to eight. The subjects in the two groups were combined and regrouped according to their score increase. Table 2 indicates the distribution of pass, fail, and not attempted scores in the first forty minutes in relation to score increase in the last twenty minutes and reveals a very striking pattern. Subjects who showed no score increase clearly have attained their upper limit of success in the first forty minutes and do not benefit by additional time. Subjects whose score increase was one, two, or three points, appear to form a homogeneous group which is not very different from the zero-increase group. However, subjects whose score increase was four, five or six points, also form a homogeneous group, but with a very different pattern of performance in the first forty minutes compared with the other groups. While their pass rate is

TABLE 2

RELATIONSHIP BETWEEN SCORE INCREASE AFTER FORTY MINUTES AND PATTERN OF SCORES DURING FIRST FORTY MINUTES (COMBINED GROUPS).

Score Increase	Score Pattern (Items 21-48) After Forty Minutes			N
	Pass	Fail	Not Attempted	
0	12·4	11·1	4·5	49
1	12·2	7·6	8·2	28
2	13·1	6·7	8·2	31
3	3·5	6·1	8·4	13
4	11·0	3·0	14·0	11
5	9·2	2·2	16·6	4
6	11·0	2·0	15·0	4
7	5·0	3·0	20·0	1
8	5·0	2·0	21·0	1

only slightly less than that of the other groups, their failure rate is very much lower and their not attempted rate very much higher. In other words, these subjects (forming about 13 per cent. of the total sample) appear to work slowly but at a high level of accuracy during the first forty minutes and, in terms of power, are severely handicapped by a time limit.* It is perhaps worth pointing out that a score increase of six points represents, in terms of the Australian norms for this test an increase in centile rank (using a score increase from 27 to 33, as an example) from 20·3 to 56·7. The intellectual level of these students will be severely underestimated if an artificial time limit is imposed on the test.

It is an open question whether this distinctive group would be handicapped or not in either their university courses or examinations by this apparent tendency to work slowly but very efficiently. In the absence of further information, either effect would be conceivable. Thus, it could be argued that they would have difficulty in completing enough high-level work under constant pressure of time (a point of particular relevance where part-time external students are concerned) or of writing enough material under examination conditions to obtain results commensurate with their ability. It might equally well be argued, however, that, although their output might be smaller than that of faster-working students, its quality would more than compensate for this.

Some information, of an inconclusive nature, is available on this point.† All the subjects involved in this study were enrolled in Psychology I and their examination results are shown in Table 3. Four categories of results were available—pass with merit ; pass ; fail (but granted a supplementary examination) ; and fail outright. The figures in brackets are percentages. It is clear that the 'high-increase' group shows a pattern of results which is fairly similar to that of the 'low-increase' group and it may be concluded that the characteristic pattern of work on the intelligence test does not appear to have any marked implications as far as passing or failing examinations is concerned. This is not surprising since success or failure in examinations is undoubtedly a function of a

TABLE 3

RELATIONSHIP BETWEEN SCORE INCREASE AND PASS/FAIL CATEGORY IN PSYCHOLOGY I EXAMINATION.

Score Increase	Examination Category			
	Pass With Merit	Pass	Provisional* Fail	Outright Fail
Low Increase ..... (0–3)	13 (10·75%)	72 (59·50%)	7 (5·79%)	29 (23·97%)
High Increase . . . (4–8)	4 (19·05%)	9 (42·86%)	2 (9·52%)	6 (28·57%)

* Candidate granted a supplementary examination

large number of factors, of which speed of working is only one. Two comments, however, may be made. Students who proceed in a cautious but accurate way are capable of succeeding (in some cases at a high level) in time-limited examinations. Secondly, it is not known to what extent students in the pass category might be capable of lifting their performance into the pass-with-merit category, given additional time ; nor is it known whether some of the students who failed the examination might have passed with more time at their disposal. These points would merit further investigation, perhaps by the use of objective-type examinations or the use of timed and untimed attainment tests.

---

* In his manual, Raven did not specify any time limit; the Australian norms published by the A.C.E.R are, however, based on a time-limit of 40 minutes.

† For information contained in this section, the author is indebted to Mr. H C. Sheath, Director of External Studies of the University of New England, and to the Heads of Departments, for kindly supplying details of examination results.

V.—REFERENCES.

FOULDS, G. A., and RAVEN, J. C. (1950). An experimental survey with progressive Matrices (1947). *Brit. J. Educ. Psychol.*, 20, 104-110.

YATES, A. J. (1961). Item analysis of progressive matrices (1947). *Brit. J. Educ. Psychol.*, 31, 152-157.

SOCIAL FACTORS AND SCHOOL INFLUENCE IN JUVENILE DELINQUENCY

AN ANALYSIS OF POLICE CASES IN THE BELFAST JUVENILE COURT, JULY, 1961, TO JUNE, 1962

BY G. A. DELL

(*Educational Psychologist, Belfast Education Authority*)*

SUMMARY. An analysis of some of the educational and social factors associated with 492 police cases appearing in the Belfast Juvenile Court over a period of twelve months, showed the heaviest incidence of delinquency among children of secondary age in non-selective schools. Differences in incidence rates between Roman Catholic and Protestant children were strongly associated with differences in socio-economic status. Other factors appearing to facilitate delinquency were a low level of socio-economic status, retardation in attainments, and possibly also particular age groups. Independently of the general socio-economic level, the location of a school in one of the old, central, and socially declining, areas of the city was associated with a high delinquency rate. The quality of a school's morale was effective in a small but measurable degree in counteracting tendencies to delinquency Ways of increasing the strength of this influence are briefly discussed.

I.—INTRODUCTION.

THE purpose of this paper is to outline some of the educational and social factors associated with juvenile delinquency in Belfast, by analysis of one year's record of police cases appearing before the Juvenile Court. The explanation of individual variation—why one child becomes delinquent and another does not—rests on psychological, constitutional, and perhaps chance factors, rather than on social data, but analysis of the social factors involved draws attention to the kind of social conditions appearing to facilitate delinquent behaviour, and therefore, to areas in which preventive measures may most economically be applied, even though these measures may, from the nature of the analysis, be compensatory rather than radical, in that they relate to the appearance of the symptom and not to its aetiology.

II.—THE DATA.

(*a*) *Tabulation of 492 Police Cases.*

These tables, derived from an analysis of school reports presented in each case, are given in full in Appendix I. In brief, they allow the following factors to be considered: type of school attended (level of education provided); type of school attended (county or voluntary administered); age distribution; level of attainment relative to type of school attended; attendance record; and differences in incidence rates between individual schools (this last table is not included in the Appendix).

(*b*) *The Distribution of the General School Population of Belfast.*

(i) The Annual Report of the City of Belfast Education Committee for 1960-1, gives enrolment figures as at 31st December, 1960. These include children attending extra-district schools, and were the latest published figures available. They have been used as the basis for comparisons between groups of schools, age groups, etc. (See Appendix II.)

---

* The views expressed are those of the author and do not necessarily represent those of the Belfast Authority.

(ii) For comparisons between particular schools, the Annual Statistical Returns for 1961 have been used. These give figures for enrolments on 12th January, 1962, classified according to age at 31st December, 1961. We are, of course, interested only in the distribution for children above 8 years of age, since these are technically the children 'at risk.'

(c) *A Measure of Socio-Economic Level of Each School.*

A crude index of the socio-economic background of each school has been obtained by using the school meals service data on the number of necessitous children receiving free meals in each school during a week chosen at random. This has been termed a 'poverty rate.' Rates were calculated for both 1961 and 1962, in both cases a sample week in May being chosen. The rates for the two years correlated highly ($Rs=0{\cdot}93$ for county secondary intermediate schools) showing the index to be a reliable one. In the main, the poverty rate for 1962 was used subsequently.

(d) *A Measure of School Quality.*

In order to try to assess the influence of school quality on juvenile delinquency, as far as possible, independent of socio-economic background, five senior officials of the Education and Health Departments were asked to rank thirteen boys' and mixed secondary intermediate schools (the equivalent of English secondary modern schools) in order of 'quality.' The five rankings were then subjected to a test of minimal significant concordance (Kendall's Statistic W) and were found to be very significantly concordant ($W=0{\cdot}80$, $P<0{\cdot}001$). It was then possible to produce a final order of quality by summing the individual rankings for each school.

III.—RESULTS.

(1) *Type of School Attended—Level of Education.*

Very highly significant differences ($\chi^2=241{\cdot}81$, $P<0{\cdot}001$) were obtained between the incidences of delinquency reported for the various types of school. The heaviest incidence lay among unreorganized primary school pupils of secondary age (1·70 per cent. of those at risk), followed by secondary intermediate school pupils (1·34 per cent.). The incidence among primary school pupils of primary age was much lower (0·32 per cent.), and that for secondary grammar schools the lowest of all (0·20 per cent.). Incidence rates for special schools were very high (1·86 per cent.) but unreliable owing to the small numbers involved. In general, both age and ability appear to be involved, as well as the possibility of other factors. The overall incidence rate for school children at risk was 0·76 per cent.

(2) *Type of School Attended—Religious and Administrative Affiliation.*

There were no significant differences between county and voluntary schools overall ($\chi^2=1{\cdot}21$, $P<0{\cdot}30$), or between county and voluntary primary school children of primary age ($\chi^2=0{\cdot}005$, $P<0{\cdot}95$), or between county and voluntary secondary intermediate schools ($\chi^2=0{\cdot}655$, $P<0{\cdot}50$) , or between county and voluntary E.S.N. special schools ($\chi^2=0{\cdot}002$, $P<0{\cdot}98$). There were highly significant differences between county and voluntary primary school children of secondary age—i.e., attending unreorganized primary schools—($\chi^2=10{\cdot}89$, $P<0{\cdot}01$) favouring county school children, and between Protestant and Roman Catholic grammar school children ($\chi^2=6{\cdot}48$, $P<0{\cdot}02$) favouring Protestant children. Even where differences were not significant, incidence rates in voluntary schools tended to be higher than those for comparable county schools (the sole exception to this tendency being in the case of E.S.N. special schools). However, when allowance is made for the socio-economic difference, and when this factor is held reasonably constant, differences in delinquency rates associated with the religious or administrative affiliation of schools, tend to disappear. For example, when county grammar schools are compared with Roman Catholic grammar schools (i.e., omitting Protestant voluntary grammar schools, which may be assumed to draw pupils from on the whole higher socio-economic backgrounds) $\chi^2=0{\cdot}15$, $P<0{\cdot}70$. From another point of view, an inspection of the poverty rates for the unreorganized primary schools, where the biggest difference between voluntary and county schools occurs, shows that the voluntary schools have a considerably larger number of necessitous children on rolls than have the county schools, size held constant. It is probable that whereas reorganization of county schools follows a plan which is relatively unaffected by the socio-economic background of the children concerned, in the case of voluntary schools the poorer parishes are understandably slower to reorganize. There is, perhaps, in consequence a selective factor operating to increase the delinquency rates of voluntary unreorganized primary schools relative to comparable county schools.

At the same time, since the poverty rates for voluntary schools in Belfast tend to be higher than those of comparable county schools, at all levels of educational provision, there is an implication that perhaps a high poverty rate may be less significant for delinquency at the primary age, and within certain ranges of ability.

(3) *The Age Factor.*

The evidence clearly shows that delinquency rates increase with age, upto the end of the period of compulsory education (i.e., at 15 years of age, in the case of Northern Ireland children). Thereafter, the rate falls again (presumably because the over-age children still at school are selected by ability, and by social class, and are less subject to delinquent influences).

There is some evidence of an abnormally large increase in the delinquency rate at 12+, and to a lesser extent another at 13+. The standard increase in incidence rate from one age group to the next appears from this data to be approximately 0·16. This is consistent up to the 12+ age group, when the increase jumps to 0·44, and for the 13+ group it increases by another 0·32. Finally, in spite of the generally and consistently increasing delinquency rate from one age group to another, the differences between the rates for 13+ and 14+ were not in fact statistically significant ($\chi^2=0{\cdot}27$, $P<0{\cdot}70$). This finding modifies to some extent the notion that the peak delinquency rate must necessarily be in the last year of compulsory schooling.

(4) *Attainments.*

There is a significantly heavier incidence of retardation among delinquents than one might expect on a chance basis ($\chi^2=104{\cdot}66$, $P<0{\cdot}001$, on the assumption that the ratios of above average : average : below average children in the school population is 20 : 60 : 20. Any other symmetrical distribution would increase χ^2 still further). For children still at school, there is no significant difference between primary and secondary intermediate schools in regard to the proportion of delinquent children who are retarded ($\chi^2=2{\cdot}87$, $P<0{\cdot}30$). However, if one includes school leavers in the assessment, the tendency towards a higher incidence of retardation among secondary school children than among primary school children, becomes statistically significant ($\chi^2=6{\cdot}74$, $P<0{\cdot}05$). This suggests that the retarded school leaver is in even greater hazard than the retarded secondary school child, in terms of delinquency. There appears to be no association between delinquency and academic standing in the grammar school, although numbers are very small. Rank order correlations between the delinquency rates for all mixed and boys' secondary intermediate schools and the rate showing the proportion of retarded children among the delinquents from each school, are consistently negative, and of the order of $-0{\cdot}30$, but do not reach an acceptable level of statistical significance. However, the consistent appearance of negative correlations here does suggest that high delinquency rate schools tend to have a generally lower standard of attainment, to the extent that the delinquent groups from these schools are less likely to be assessed as retarded.

(5) *Attendance.*

The tendency appears to be for delinquents as a group to show somewhat poorer attendance records than one might expect from non-delinquent children, but there is no adequate basis of comparison here, and many delinquents seem to have been excellent attenders, and the great majority satisfactory in their attendance. There is, in fact, no evidence for supposing that regularity of attendance in itself has any special bearing on delinquency.

(6) *Socio-Economic Status of School Background.*

Correlations between this factor and delinquency rates are fairly consistently positive, high, and statistically significant, when groups are analysed that are reasonably homogeneous by sex and religious affiliation. For example, for county secondary intermediate schools—excluding girls' schools—Rs is $0{\cdot}88$, $P<0{\cdot}01$. For all voluntary secondary intermediate schools Rs is $0{\cdot}79$, $P<0{\cdot}05$. For both county and voluntary secondary intermediate schools—excluding girls' schools—taken together Rs drops to $0{\cdot}65$, $P<0{\cdot}05$. For eighteen county primary schools having delinquents Rs is $0{\cdot}46$, $P<0{\cdot}05$.

A strong association between high poverty rates and high delinquency rates is clearly established, especially at the secondary intermediate school level.

(7) *Size of School.*

Rather surprisingly, significant associations have been found between small schools and high delinquency rates. One might have expected to find that the more personal atmosphere of a small school would be associated with lower delinquency rates, but, in fact, the correlation between size and delinquency rate is a negative one at both the primary and the secondary level. For county primary schools having delinquents—eighteen schools in all—Rs is $-0{\cdot}76$, $P<0{\cdot}01$. For county secondary intermediate schools—ten boys' and mixed schools—Rs is $-0{\cdot}68$, $P<0{\cdot}05$. At the primary school level one might attribute part of this correlation to the fact that the calculation of delinquency rates from small numbers of delinquents tends to favour the larger primary schools. However, the same objection cannot be made to the results from secondary schools, where no school has less than 500 children at risk, and yet the correlations are roughly of the same magnitude and in the same direction.

Correlations between 'size of school' and 'poverty rates' are also of interest. For the eighteen county primary schools Rs is +0·09, N.S. For the ten county secondary intermediate schools Rs is – 0·83, P < 0·01. The small primary schools are thus not 'impoverished,' whereas the smaller secondary intermediate schools do tend to be.

The rationale behind this particular group of findings may be as follows :

(*a*) The larger secondary schools tend to be the newer schools, and schools in newer housing areas, where the general socio-economic level may be expected to be of a higher order than in the older areas of the city served by smaller schools.

(*b*) While there may be no significant association between size of primary school and socio-economic level, inspection of the list of schools concerned shows that a number of the smaller primary schools are high up in the delinquency list and also in the poverty ranking. It is also apparent that these schools tend to be congregated in certain areas of the city—notably in Court Ward, St. Anne's, and other central areas. Their common characteristic may be not merely a low level of socio-economic status, which may be true of many other primary schools in Belfast, but also being centred in what E. Jones (1960) has described as areas of declining importance from the social and residential point of view, where space is at a premium, and where the general trend of migration is outwards for the more ambitious and inwards for the socially deprived.

In the case of the county intermediate schools, while socio-economic status clearly must contribute to the size of the correlation between size of school and delinquency, when socio-economic status is partialled out, using Kendall's Tau and partial rank correlation technique (Siegel 1956), there is still a substantial negative correlation between size and delinquency rate (Tau= –0·43, N=10) which may well be statistically significant.

(8) *School Quality.*

A. B. Clegg (1962d) has reported an investigation with thirty West Riding of Yorkshire secondary schools, in which he found a correlation of 0·50 between school quality (as measured by rankings of a panel of four senior officers of the Education Department) and Delinquency rate, with high quality schools tending to show lower delinquency rates. He further found "no correlation at all" between School Quality and the Social Desirability of the areas in which the schools were located (as measured by the rankings of senior county council officials), and a correlation of 0·02 between Delinquency and Social Desirability. He concluded that "the evidence from these schools supports the commonsense view that, even in adverse circumstances, the school itself can do much to combat delinquency and juvenile crime." For thirteen Belfast boys' and mixed secondary intermediate schools this analysis shows a correlation between low quality and high delinquency rate well in line with Clegg's finding (Rs=0·56, P<0·05). But contrary to Clegg's finding, the correlation between socio-economic status, or 'poverty rate,' and delinquency rate, is even higher (Rs=0·65, P<0·05). The correlation between school quality and poverty rate is 0·63, P<0·05.

In fact, if socio-economic status, or 'poverty rate,' is partialled out (now using Kendall's tau throughout) the correlation between school quality and delinquency rate drops from Tau=0·45, P<0·05 to Tau=0·29 (no tests of significance available, Siegel 1956).

Thus, Clegg's conclusion is not so optimistically borne out by this analysis. There is certainly a measurable effect due apparently to school quality, or to some related factor, but this effect seems to be considerably less than that of socio-economic influences. Indeed, if this were not so, any school with a high delinquency rate would be saddled with some burden of guilt for failing to combat the seeds of crime.

IV.—LIMITATIONS OF THE DATA.

(*a*) Possible errors in the data: duplication due to recidivism has been eliminated (reducing the number of cases from 508 to 492), and cases who had already left school at the time of their court appearances have also been identified (leaving 418 cases of children still at school). However, other minor classificatory errors may still exist—e.g., a child before the court may be the subject of a report from a school which he, in fact, no longer attends, or did not attend when he committed the delinquent act.

(*b*) The data must, of course, be regarded as applying to 'police cases before the Juvenile Courts' rather than to Juvenile Delinquency in Belfast in general. Intelligent delinquents, delinquents from well-to-do homes, and younger delinquents may well be under represented.

(*c*) No distinction has been made between male and female delinquents, except incidentally when this is revealed through the schools attended. The universal finding is that male delinquency greatly exceeds female delinquency, particularly during school ages. For example, in 1954, of 337 children appearing before the Belfast Juvenile Court in police cases, only sixteen were girls. It is not unreasonable, therefore, to exclude girls' schools from this analysis, which is what in fact has been done. However, by the same argument we should strictly speaking consider only the boys at risk in for example mixed secondary schools, otherwise a comparison including both boys' and mixed secondary schools will weight the delinquency rates of the boys' schools unfairly. In fact, although only correlations based on delinquency rates calculated from total school enrolments have been quoted in Section III, similar data was also obtained for delinquency rates calculated for boys' enrolments only, and this reproduced similar patterns of results in terms of the size and direction of the correlation coefficients.

(*d*) No distinction is made between indictable and non-indictable offences, and all police cases are assumed to be of equal 'social significance,' although the typical delinquent act may, in fact, vary from age group to age group. However, at all the ages under consideration, the great majority of offences relate to some form of Larceny and/or Breaking and Entering, so that this objection is not particularly powerful in this context.

(*e*) For some of the data there is no adequate standard of comparison—for example, 'level of attainient' is a purely relative and probably rather subjective assessment. Similarly, for 'level of school attendance' there is no adequate control data.

(*f*) The index of socio-economic background used here is a very partial estimate, stressing only one aspect of status, and may possibly introduce an element of age selection—e.g., families with older children may be less likely to fall within the 'necessitous' group. Also, a status index of this kind can be misleading in that it is not a continuous measure of status from low to high, but an index only of the amount of low status in any given school.

(*g*) The measures of ' School Quality ' and ' Socio-Economic Status ' are apparently highly reliable, but they may not necessarily be valid measures. It is possible for example, that the ' School Quality ' index may be in part just a ' better ' measure of ' Socio-Economic Status ' than is the Poverty Rate used for that purpose.

V.—DISCUSSION AND CONCLUSIONS.

The general pattern of these results is one showing the heaviest incidence of delinquency among children of secondary age in non-selective schools. Within the non-selective secondary schools a number of factors appear to facilitate or to be associated with delinquency—notably, a low level of socio-economic status, retardation in attainments, an age factor, and the possibility that particular age groups (e.g., 12+) may be specially susceptible. Independently of the general socio-economic level, the location of a school in one of the old, socially declining areas of the city seems to be associated with an increased delinquency rate, and the quality of a schools' morale or influence seems to be effective only to a rather limited degree in overcoming and counteracting more powerful tendencies to delinquency.

Typically then it is not the near-failures at 11+ who tend to become delinquent in the main (in the secondary intermediate schools with good academic standing and opportunities, the proportion of retarded children in the delinquent group is higher than in those schools of generally lower academic status). It is rather those who are academically retarded, or who lack academic goals, who are most susceptible. The selection process itself does not seem to be a significant factor, although the educational segregation that follows it may tend to minimize the compensatory influence that ' school quality ' might otherwise exert.

In general, these findings confirm those reported for similar areas in the United Kingdom. This is in itself of interest, since the consistently high level of unemployment in Belfast, and the existence of deep-rooted religious and social differences which have in some respects produced parallel cultures within the same geographical community, might have been calculated to exert a differentiating influence on delinquency rates here. Barritt and Carter (1962) have shown, however, that socio-economic differences may account for a great deal of what ostensibly appears to be difference due to religious affiliation, because socio-economic status in Northern Ireland is correlated with religious affiliation. The evidence here suggests that the same is true of apparent differences in delinquency rates as between Roman Catholic and Protestant schools, where these are statistically significant.

The influences of socio-economic status and educational retardation have been attested by among others, Ferguson (1952) on Glasgow data, and Burt (1944) in his analysis of London children after the first World War, concluded that while poverty in itself made a comparatively small contribution to delinquency, conditions associated with poverty, or facilitated by it, were of the very highest significance (e.g., defective discipline especially in large families, intellectual disabilities, detrimental interests, neighbourhood influences, etc.). But while the influence of this ' poverty complex ' appears to be strong in the Belfast data, and in spite of the high unemployment rate in Belfast over many years, nevertheless, delinquency rates in Belfast compare not unfavourably with those of most other large cities in the United Kingdom. The ' poverty complex ' operates within a cultural setting in which other factors can apparently modify its influence considerably.

The particular modifying influence on which this analysis provides some evidence, is that of the school. It is commonly assumed that the school can be a positive force against delinquency. The Scottish Advisory Council on Education (1952) concluded that the evidence " gives some encouragement to teachers who believe that the school can exert a favourable influence on pupils living in an area where delinquency is common." Mays (1962) argues that reducing the incidence of juvenile delinquency is basically an educational problem. Clegg (1962 a, b, c, d) in a series of brief reports and memoranda, has written most stimulatingly on this possibility, although he describes the attitude of some members of the teaching profession to this question as one of reserve, commenting that " it is almost as if they were prepared to forgo the credit for inhibiting delinquency if they can escape censure for not doing so." (Clegg, 1962 d). However, this analysis of Belfast data seems to indicate that the influence of ' school quality ' is much less than that of say ' poverty.' How can one reconcile the optimistic and the less optimistic results ? Under what sort of conditions may ' school quality ' become a powerful influence in combatting delinquency ?

On the first point there are, of course, many possible explanations of this discrepancy—possible errors of measurement and analysis aside, there are differences in the overall level of delinquency rates, and probably in genera socio-economic level as measured by a ' poverty rate ' ; the West Riding data refers only to indictable offences, and samples a wide area using broad assessments of ' social desirability,' while the Belfast analysis covers a small area, using a narrow but more finely graded measure of ' poverty.' Any or all of these factors could conceivably tend to affect the apparent influence of school quality as against socio-economic background. Similarly, the inclusion or exclusion of girls' secondary schools could have a very large effect on the size of the correlation between school quality and delinquency, and between socio-economic background and both these factors. Finally, it is possible that the West Riding estimates of ' school quality ' are less contaminated by judgments based partly on socio-economic impressions, or indeed, that the West Riding schools may have been more successful in influencing delinquency rates—both for better and for worse. But the similarity between the two findings in regard to the correlation between Delinquency Rates and School Quality (West Riding 0·50, Belfast 0·56) makes this seem improbable, and the total absence of correlation between School quality and Socio-economic background in the West Riding data itself requires explanation. It seems not unreasonable, therefore, to say that at least under the general prevailing social conditions in Belfast, the influence that the school can exert on delinquency rates should not be exaggerated.

On the second point, we enter the realms of speculation, and as Clegg (1962 b) points out we urgently need studies of ' good ' and ' bad ' schools from the standpoint of delinquency. But from what facts we have it seems that if the quality of a school's morale can influence delinquency rates as the evidence shows that to some degree it already does, then the real problem is how to manifest that quality in the lower streams equally with the top. In recent years, the stress on quality and parity has been related to providing opportunities for the 11+ failures. If we ranked our schools for the quality of their bottom streams we might find a much higher correlation with delinquency rates. In larger secondary schools of good academic standing, the organization of departments for the retarded streams, with some autonomy within a flexible framework, seems promising. In smaller, or less strong academically, secondary schools a more experimental approach is urgently required. In some cases this might mean experiment with an unstreamed organization (Rowe 1959). Clegg (1962, a, b) has described very clearly problems and methods in the development of school discipline. Equally important is the need to develop avenues of legitimate achievement, for the solitary as well as for the gregarious. Priority for these schools and areas in the development of Youth Centres, and Further Education Centres, in allocations of time, space, and money for games and other shared facilities, in the institution of counselling relationships between teachers and pupils that can be sustained beyond the period of school attendance, in the arrangement of small group school visits rather than the more typical large group visits—these may be some of the ways in which relief can be given, and response obtained.

Perhaps the first priority might well be for an education authority to adopt its school with the consistently highest delinquency rate, as an experimental school, and to try to see how organization, curriculum, activities, relationships, and the supporting facilities of the surrounding community, could be developed by a joint effort to make an inroad against delinquency.

VII.—References.

Barritt, D. P., and Carter, C. F. (1962). *The Northern Ireland Problem.* London O.U.P.

Belfast Education Committee (1961). *Annual Report for* 1960-1. Belfast Education Committee.

Burt, C. (1944). *The Young Delinquent.* 4th Edn. London: U L P

Clegg, A. B. (1962 a) *Caning, Behaviour, and Delinquency in Secondary Schools.* W. Riding C.C. Education Committee

Clegg, A B. (1962 b). Delinquency and discipline: 4—The role of the school *Education*, 22nd June, 1962.

Clegg, A. B. (1962 c) *Juvenile Crime in the Schools* Private memorandum.

Clegg, A. B. (1962 d). Social factors *Education*, 22nd June, 1962.

Ferguson, T. (1952). *The Young Delinquent in His Social Setting. A Glasgow Study.* London: O.U P

Jones, E. (1960). *The Social Geography of Belfast* London: O U P.

Mays, J. B. (1962). Delinquency and discipline 3—The influence of environment. *Education*, 15th June, 1962.

Rowe, A. W. (1959) *The Education of the Average Child* London: Geo G. Harrap.

Scottish Advisory Council (1952). *Pupils who are Maladjusted because of Social Handicaps.* Edinburgh: H M.S.O.

Siegel, S. (1956). *Non-Parametric Statistics* New York: McGraw-Hill

VIII.—Appendices.

Appendix I.—Tabulation of 492 Police Cases.

TABLE 1A

Type of School Attended.

	County Schools			Voluntary Schools			
	At School	Left School	Total	At School	Left School	Total	Total
Primary—Primary age	44	—	44	29	—	29	73
Primary—Secondary Age	20	12	32	57	16	73	105
Secondary (Intermediate)†	160	27	187	74	18	92	279
Grammar	4	1	5	15*	—	15*	20
Special (E.S.N.)	10	—	10	5	—	5	15
Totals	238	40	278	180	34	214	492

*—Includes 4 boys attending a voluntary non-R C. grammar school.
†—'Secondary Intermediate' is equivalent to 'secondary modern.'

TABLE 1B

AGE RANGE OF DELINQUENTS

Age Groups	8+	9+	10+	11+	12+	13+	14+	15+	16+	Total
Attending School..	4	15	26	37	70	99	109	28	30	418
Left School ..	—	—	—	—	—	—	—	44	30	74
TOTALS ..	4	15	26	37	70	99	109	72	60	492

TABLE 1C

ATTAINMENTS OF DELINQUENTS.

	PRIMARY		SECONDARY		GRAMMAR		Total
	At School	Left School	At School	Left School	At School	Left School	
Above Average	8	1	12	2	2	0	25
Average	92	18	124	15	14	1	264
Below Average	50	9	98	28	3	0	188
	150	28	284	45	19	1	477

TABLE 1D

ATTENDANCE OF DELINQUENTS

	At School	Left School	Total
Satisfactory/Excellent	82·1%	77·2%	81·1%
Irregular/Bad	17·9%	22·8%	18·9%
TOTAL NUMBERS INVOLVED	418	74	492

APPENDIX 2.—Distribution of School Population of Belfast in Relation to Delinquents.

TABLE 2A

TYPE OF SCHOOL ATTENDED

	County Schools		Voluntary Schools		Total Numbers	
	D	P	D	P	D	P
PRIMARY—Primary age ..	60·3%	60·7%	39·7%	39·3%	73	22812
PRIMARY—Secondary age	26·0%	49·8%	74·0%	50·2%	77	4532
SECONDARY (Intermediate)	68·4%	71·0%	31·6%	29·0%	234	17444
GRAMMAR	42·2%*	73·1%*	57·8%	26·9%	19	9674
SPECIAL (E.S.N. only) ..	66·7%	66·2%	33·3%	33·8%	15	535
* Includes voluntary non-R.C schools D=Delinquents. P=Population			Add (Other 'Special')			.. 273
			Totals		418	55270

TABLE 2B

AGE RANGE

Age Groups	8+	9+	10+	11+	12+	13+	14+	15+	16+	Total No.
% of school population aged 16+ .	12·1	11·9	12·0	12·3	12·7	13·5	13·1	12·4		55270
% of delinquents at school	1·0	3·6	6·2	8·8	16·7	23·7	26·1	6·7	7·2	418

COMMENTS ON

SOCIAL FACTORS AND SCHOOL INFLUENCE IN JUVENILE DELINQUENCY BY G. A. DELL.

BY A. B. CLEGG

(*The Education Officer, West Riding of Yorkshire County Council*)

Mr. Dell's investigation is of considerable interest and confirms the results of Burt and others in approximately similar urban environments. The main difference between his results and the results of what was a very simple investigation by a local authority in the West Riding lies in the relationship between school quality and socio-economic status. I am satisfied from the West Riding enquiry that a school in which the morale is high because the children are finding school life worth while can, and indeed often does, make a contribution to the standards of behaviour in the area. Whether this could happen in a dense urban area where standards were really low, is another matter. I personally believe it could and that in some urban areas evidence could be found of its having done so. It would, of course, be an interesting and valuable piece of pioneer work if, as Mr. Dell suggests, a large Authority could staff a school in a really bad area with a task force of skilled and sensitive teachers and then follow up the results to see whether they had made an impact on the wider social environment.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

A Projective Technique for the Assessment of Occupational Interests, with Special Reference to Creative and Artistic Capacities and Decision Making.

By LISBETH BRIESS, M Sc., Ph.D.

(Abstract of a Thesis submitted for the M.Sc. degree in Psychology of the University of London.)

AIM.

To devise a projective technique, capable of being scored objectively, for the assessment of occupational interests.

THE TECHNIQUE.

The Subjects (S) are presented one by one with 75 ambiguous drawings on cards, S being permitted only one of two given interpretations ; S has to decide whether, for instance, a drawing represents a human being or an animal, a human being or a slab of stone, a dog or a car, in a way similar to that used by Frenkel-Brunswik[2] and Vernon Hamilton,[3] but so designed that perseveration is excluded. There are 25 cards with the alternatives ' human or animate ' (' animate ' means animal or plant), 25 cards with the alternatives ' animate or inanimate,' and 25 cards with the alternatives ' human or inanimate ' If S is unable to make a choice he has to state that he cannot decide and that it could be either. Another 25 cards with unstructured drawings are mixed amongst these, and S is asked whether the drawings are meaningful or not meaningful for him.

The 75 cards with alternative choice have the purpose of assessing (*a*) a Subject's basic interests ; (*b*) his distastes, the importance of which has been stressed by Rodger[4] ; and (*c*) his general ability to make decisions. The 25 unstructured drawings give information about his capacity for doing creative and/or artistic work.

Every drawing had to satisfy a criterion before it was incorporated in the set for the actual experiment. The picture had at least once in eight successive responses to be interpreted in both ways and this had to happen for three consecutive sequences of eight interpretations. It was ascertained that there were no significant differences in the interpretations between men and women, and between Europeans and non-Europeans who had lived in our society for some time

After an unscored trial with three other pictures, the thoroughly shuffled cards are presented one by one, and S is instructed to give the very first impression as quickly as possible, but there is no time limit set. Questions are then asked about occupation, interests and leisure activities. The scores consist of the number of choices S has made for every one of the five categories, called briefly ' Human (H),' ' Animate (A),' ' Inanimate (I),' ' Either (E),' and ' Not meaningful (NM).'

After a preliminary pilot experiment two experiments were carried out in a very similar way, but with different sets of cards in order to make sure that the results were not attributable to features of the cards

EXPERIMENT I.

This experiment was carried out on fifty Subjects. They consisted of members of the academic, administrative and technical staffs and students of several departments of Birkbeck College and the London School of Economics, and social workers. Their ages were 17 years to the middle 50s., and there were about as many men as women.

Reaction-times were taken between the Experimenter reading the description of a card and S commencing to say his choice This experiment gave highly significant results in support of the hypotheses that a Subject's general ability to make decisions and his capacity for doing creative and/or artistic work can be assessed in this way, if his answers to the questions about special interests and spare-time activities may be taken as a crude criterion prior to a systematic validation. The results of this experiment with regard to interests are, as likely as not, due only to chance, as the sample was too heterogeneous. No attempt was made with this heterogeneous sample to investigate distastes.

The mean E-scores for those S's who knew what kind of occupation they wanted were significantly lower than for those who had not yet made up their minds. As older people, who are already in occupations, are trained at Birkbeck College, the sample was very useful for this part of the investigation. The Mann-Whitney test[5] for the individual E-scores gave a significant result at the 1 per cent. level. The reaction-times before scoring an E were significantly higher than before scoring any other category. By analogy with the claims for word-association tests, the assumption cannot be excluded that the inability to make a decision has such emotional repercussions, even in a simple experiment, that there is a correlation between underlying emotions and reaction-times.[4]

The S's with no artistic /creative interest or spare-time activities had signicantly higher NM means than the others.

EXPERIMENT II.

This experiment was carried out in a manner similar to Experiment I, but with specially selected Subjects and a new set of drawings. Two samples were obtained, one mainly of thirty-three students of science and engineering, and junior laboratory technicians who intended to become university students ; the second mainly of forty-seven casework students and students who had the special desire to become teachers. The S's were all Europeans, 17-25 years old, with at least qualifications for entering a university. They were not aware that they had been selected for the sample as social workers, scientists, etc These samples yielded sufficiently encouraging results with regard to interests to make the investigation of other occupations worth while, although the validation has to be repeated because the samples were small. Reaction-times were not taken with this sample to make sure that there was no distraction from the process of externalisation.

With this sample, distastes were investigated in the following way. If S scores much lower in one or two of the categories H, A, or I than in the others, this item may concern S so little, or be so strongly repressed that any dwelling on it would produce a distaste.[6] It has been postulated that a distaste is indicated if Chi-square between the lowest and the highest scores is significant at the 5 per cent. level. Distastes according to this postulate for one or the other category of interests are indicated with high frequency.

FACTOR ANALYSIS.

With the data of Experiment II Stephenson's inverted technique of factor analysis,[7] by which Subjects are correlated, was performed, and multi-factor analysis, as described by Cattell,[1] yielded four general factors which coincided with personality dimensions.

The results of one factor analysis have, of course, to be interpreted with caution, But since the M.Sc. examination, the factor analysis has been repeated on two more samples, and the same patterns have emerged for the loadings of the four factors ; thus, the procedure has yielded the same factors three times. One of them is a bipolar factor for what Super[8] has described as " . . . scientific interest, technical or material interest, humanistic or social welfare interest . . . ", which contrasts scientific interests on one side with humanistic or social welfare interests on the other.

REFERENCES.

[1] CATTELL, R. B. (1952) *Factor Analysis*. New York : Harper and Brothers.

[2] FRENKEL-BRUNSWIK, E. (1949). Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Pers.*, 18, 108-43.

[3] HAMILTON, V. J. (1957). Perceptual and personality dynamics in reaction to ambiguity. *Brit. J Psychol.*, 48, 200-215

[4] JUNG, C. G. (1960). *Studies in Word Association*. London : Routledge and Kegan Paul.

[5] MANN, H. B , and WHITNEY, D. R. (1947). On a test whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.*, 18, 50-60.

[6] RODGER, A. (1959). How should we think about ' Interests ' in occupational Psychology? *Bull. de L'Assoc. Internat. de Psychol. Appliquée*, 8, 3-9.
[7] STEPHENSON, W. (1936). The inverted factor technique. *Brit. J. Psychol*, 26, 344-361.
[8] STRONG, E K. (Jr.) (1943). *Vocational Interests of Men and Women.* Stanford University Press.
[9] SUPER, D. E. (1957). *The Psychology of Careers.* New York : Harper and Brothers.

A Survey of the Attainments and Interests of Ten-year-old Children from Differing Geographical Environments within a County.

BY A. N. SMITH

(*Abstract of a thesis for the degree of M.Ed., in the University of Manchester,* 1962.)

AIMS.

The survey aimed to discover whether the attainments of 10-year-old children in English and Arithmetic and their interests, differed according to their environments : rural, urban and ' fringe.'

Subsidiary questions were :

(*a*) Is there a uniform pattern of interests amongst the 10-year old children in different localities ?

(*b*) Do children's academic and social advantages vary according to geographical environments ?

THE CHILDREN STUDIED.

The survey was limited to twelve maintained schools, chosen largely because their head teachers agreed to co-operate in the investigation, but also distributed over different geographical environments within a county. The schools were then classified, with the assistance of their head teachers, as ' urban,' ' fringe ' or ' rural.' All the 10-year-old children in the ' fringe ' and ' rural ' schools were studied, but in the ' urban ' schools a sample was taken from an alphabetical list. This was a 50 per cent. sample selected by taking every other name on the list.

METHOD.

Three main areas of environmental influence upon the children were examined—the school, the home and the neighbourhood Information about all three areas was collected by guided-interview techniques using a questionnaire guide, and also school log books, census reports, school records and personal observation. All interviews were conducted by one person to ensure a set pattern of procedure. Attainment scores were extracted from primary school record cards and a modified form of an interest schedule was also used (Blyth, W. A. L., 1954).

ANALYSIS OF DATA

(1) *The School Environment.* The schools of the urban, fringe and rural localities were compared in a number of respects ; the number, training, qualifications, experience and interests of the teaching staff ; the physical conditions and history of the schools ; the school children ; the organisation and the curriculum of the schools ; the contact between children and teachers and between the school and the home.

(2) *The Home Environment.* The occupational groupings of the parents, the education of the parents, the size of the family, family interests and contacts and the use of the mass media in the family group, were all compared between one locality and another

3) *The Neighbourhood Environment.* Comparisons were made noting the facilities in the homes of the locality, the size and nature of the gardens, the numbers of playmates, the play areas to be found and the availability of transport.

(4) *Attainments.* The scores of Moray House English, Arithmetic and Verbal Reasoning tests for 9-year-old children were extracted from the School Record Cards and comparisons between the three types of environment were made. The scores of Southend and Vernon Reading and Arithmetic tests were also extracted from School Record Cards and compared.

RESULTS.

The fringe area children appeared to have more advantages which stemmed from the home. The correlation between localities and the occupational groupings of the fathers was significant at the 1 per cent. level, but some of the advantages appeared to originate from locality rather than socio-economic sources. Thus, the lack of suitable play areas for urban children was a result of the nature of the urban locality, as were the limited contacts with other children in the rural areas.

In an attempt to derive an overall impression of the environment, the urban, fringe and rural areas were assessed as being above-average, average or below-average in respect of the many aspects examined and ' plus ' or ' minus ' scores were calculated. These scores summarise the advantages or disadvantages of the three areas.

Locality Division	' Plus ' Scores	' Minus ' Scores	Ratio of ' Plus ' to ' Minus ' Scores
Urban	7	18	1 : 2·6
Fringe	16	2	8 : 1
Rural	5	11	1 : 2·2

The cofficient of contingency for the above figures is significant at the 1 per cent. level.

CONCLUSIONS.

(1) The children from different geographical environments are influenced in different ways. (The differences may be diminishing because of improved communications and the influence of the mass media.)

(2) The urban children appeared to have the greatest number of unfavourable environmental conditions (more limited background, fewer amenities in the home, less adequate play facilities) although they have some environmental advantages (easy contact with other people, entertainment facilities readily accessible).

(3) The children of the ' fringe ' areas have considerably more advantageous environmental conditions, the limitations of the area being easily outweighed by its advantages.

(4) Children in the ' fringe ' areas showed higher attainments in English and Arithmetic than those in either the urban or the rural area, between whom there were no significant differences in this respect. These advantages and those referred to in conclusion (3) above, may have been associated as much with the socio-economic status of the parents as with the nature of the locality.

(5) The interest schedule suggested that there were some locality differences in the interests of the children, although it was again probable that the mass media and improved communications have minimised the differences. There was considerable overlap of interests between the areas

(6) Thus, within its sampling limits, this survey indicates the existence of relatively few locality differences which are not explicable in socio-economic terms

REFERENCE.

BLYTH, W. A. L. (1954) An Exploratory Enquiry into the Relation between Habitat and Outlook among Children aged Ten and Eleven. University of Leeds, *Researches and Studies*, No. 9, Jan.

An Investigation into the Efficiency of Schools Television Lessons in the Teaching of Elementary Science.

By ALBERT ROYDS

(*Summary of thesis presented for the degree of Ph.D. in the University of Manchester,* 1963.)

AIM.

The hypotheses set up for investigation were :

(1) That, after viewing a schools television lesson in elementary science, pupils' recall of its subject matter is influenced by :
  (i) factors in the structure of the lesson and its presentation, such as : lesson content ; lesson method ; and programme structure ;
  (ii) factors in the pupils' own natures and environments, such as : their sex ; intelligence ; geographical distribution , type of school ; their socio-economic grade.

(2) That there is interaction between the factors of (i) and (ii) mentioned above, i.e., between the lesson and the pupils' natural and environmental characteristics.

METHOD.

1,578 pupils, 872 in primary, 649 in secondary modern, and 57 in grammar schools, wrote answers to questions based on the content (subject matter and presentation details) of an elementary science lesson, which they viewed under classroom conditions, as part of a complete series of television lessons.

RESULTS.

The principal results are summarised as follows :

(1) The factor which has the greatest influence on recall of subject matter is the grade of pupils' intelligence. Their intelligence grade is also shown to influence their intention to attempt post-programme activities suggested for them.
(2) Girls evince greater interest than boys.
(3) Pupils' age has an effect on ; their interest, their recall of vocabulary, of scientific applications, and of attention-gaining devices from the lesson programmes.
(4) Urban pupils show greater interest than rural. Urban pupils also show better recall of basic facts.
(5) Factual learning of science applications is greater for secondary than for primary pupils.
(6) Pupils in the upper social grade show greater interest and greater intention to undertake suggested activites.
(7) Pupils' interest in the lesson programme topics shows a relationship with teachers' estimate of their teaching importance.
(8) In the presentation of the lessons, matters which influence pupils' recall include :
  (i) topics which involve reasoning, either by the narrator or by the pupil, are more effective than those stated dogmatically ;
  (ii) repetition helps pupils' interest and their intention to undertake activities ;
  (iii) length of time devoted to a topic, its position early in a progamme, and its narration at conversational pace, add to the impression made ;
  (iv) appearance of reality, association with living (including human) creatures, and mobility, contribute to effective recall.
(9) Pupils learn vocabulary from the lesson programmes.

CONCLUSIONS.

The results of the investigation support the experimental hypotheses.

BOOK REVIEWS

BIGGS, J. B. (1962). *Anxiety, Motivation and Primary School Mathematics.* London: National Foundation for Educational Research. Occasional Publication, No. 7, pp. 388, 30s.

This published thesis of an extensive inquiry into children's achievement in arithmetic, in relation to anxieties and to teaching method, is an interesting contribution to the growing literature on children's learning and achievement in mathematics. It is a publication for research workers for its detailed and specialised treatment make it unsuitable reading for readers lacking knowledge of psychology and psychometrics , but the problems it raises, perhaps more than the conclusions arrived at, are of universal interest to teachers of mathematics.

Dr. Biggs briefly, but comprehensively, considers research into the nature of mathematical ability, and observes that whereas in theories of learning one basic process is postulated in the theories of Wertheimer, Piaget, and Dienes, different kinds of thinking are distinguished. In particular he contrasts learning and problem solving which result from appreciation of structure with that which takes place through rote learning or some procedure which does not involve insight into structure ; and he questions whether they can be considered as different aspects of one process.

In discussing the role of anxiety in learning, he finds both contradictions and inadequacies in existing theories. He is critical of the equation of drive with anxiety, but disagrees with the finding that anxiety is always interfering. Biggs postulates that those easily aroused to anxiety will perform better until the difficulty of the task causes anxiety to mount to an inhibiting degree, and that inhibition is less likely to set in if the method of learning is insightful. He sees anxiety in relation to arithmetic as resulting from innate tendency to anxiety, to methods which employ anxiety as a stimulus and to the difficulty of the task.

For this major investigation, involving nearly five-thousand children from eighty-seven schools, teaching methods were classified as traditional, structional or motivational. (In traditional methods responsibility is not delegated to the children ; they are taught as a class, use a text-book and are rarely allowed to co-operate. In contrast, structural materials are intended to be used by individuals so that mathematical relations may be discovered at the child's own pace. Motivational methods use centres of interest, often miniature life situations to encourage insight into the use of numbers and measures and children often co-operate in their work). The tests used were to assess intelligence, four aspects of anxiety, and three kinds of arithmetical achievement—mechanical arithmetic, problems, and understanding of concepts. In addition, ratings of social status, parental interest and ability were made for each child by teachers. Use of the tests raises problems : girls gain higher scores than boys in general anxiety, but this may be because they are more willing to admit anxiety or because boys feel a need to falsify responses ; concept tests cease to test concepts when the relationships are easy or well known—the response is changed from an insightful one to an automatic one ; in tests of mechanical and problem arithmetic traditionally taught children are already test sophisticated whereas motivationally taught children are rarely, if ever, tested. An unexpected difficulty to the reader is that the concept tests specially prepared for this investigation are not included in the text

Statistically, the results are dealt with by a factor analysis and analyses of variance. The chief factors which appear are of arithmetical progress, anxiety attitudes centred around school, and anxiety related with mechanical arithmetic. and formality of teaching. It is somewhat tedious to follow no less than a dozen exploratory three-way analyses of variance, but they provide numerous significant interactions which deserve further investigation. Among the more interesting indications are that methods of teaching which emphasise insight induce less anxiety; that performance, especially in mechanical arithmetic, worsens steadily as number anxiety increases (or *vice versa*) ; 'and that girls tend to favour rote learning and trial and error methods whereas boys show more appreciation of structure. An additional smaller investigation suggests that where girls have been taught for several years by methods giving abstractive experience, they use insightful methods more readily in working problems.

This is an investigation which should stimulate other workers in the field in attempts to understand how children learn to think mathematically, what motivates them and in what ways insightful thinking can be encouraged. R. M. BEARD.

COOLEY, WILLIAM W. and LOHNES, PAUL R. (1962). *Mathematical Procedures for the Behavioral Sciences.* New York: John Wiley and Sons, Inc., pp. X+211.

This book presents computer programmes in the Fortran code for multiple and canonical correlations, two and three way analysis of variance and covariance, multiple discriminant analysis, principle components, factor scores and varimax orthogonal rotated factors There are still far too few programmes in this country for the main statistical procedures used by psychologists, and this volume will, no doubt, interest the expert programmer. The snag is that, although coding on difference machines is largely based on common principles, these are less important for practical pruposes than the idiosyncracies of individual machines. Few people capable of understanding Fortran will need it to programme their Pegasuses, Mercurys and Atlases. On the other hand, the flow charts for programming any digital computer should prove very useful. F. W. WARBURTON.

FOSHAY, A. W., THORNDIKE, R. L., HOTYAT, F., PIDGEON, D. A., and WALKER, D. A. (1962). *Educational Achievements of Thirteen-year-olds in Twelve Countries.* Hamburg: Unesco Institute for Education, pp. 68, n.p.

This book is based on the sound philosophy that comparative education could well utilise some methods employed in current psychological research. Whether the corresponding practical evaluation of that philosophy is the most suitable for its task is a matter of considerable conjecture, for any *measurement* evaluation as between countries, must take careful account of differing underlying principles and educational assumptions so that any statistical treatment demonstrates rather than suppresses the differences involved Tests of attainment are particularly susceptible to national differences in curricula, philosophies, teaching methods, and a variety of individual attitudes Moreover, the susceptibility is to the whole profile of such influences, and no one influence is necessarily more important than another.

The present study breaks new ground in organising and reporting on a survey of attainments in four school subjects (Mathematics, Reading Comprehension, Science, Geography), and on the variations in ability ranges as shown by a non-verbal intelligence test taken by groups of thirteen-year-olds in twelve countries. The methods of sample selection and the relative sizes varied somewhat from one country to another, and the authors acknowledge this as part of the exploratory nature of their work. The main comparisons between countries are by means and standard deviations treated in a variety of ways, and these statistics suggest that Reading Compehension shows least difference with Mathematics showing the greatest. A tentative study by Hotyat of profiles of test results for two groups of three countries each, is extremely interesting, and worthy of much cogitation for future development. Likewise, Walker's study of teacher reactions is vital to an investigation of this nature. Other evidence includes such aspects as sex differences, relationships of scores to parental occupation levels, and information on the variation of item-difficulties from one country to another.

Certain problems seem to obscure the conclusions First, any attainment test particularly, in Mathematics or Science) has a close relation to the syllabus covered and this latter is known to differ considerably between England and Europe, as also, between other countries. Differences in general teaching approaches must also be taken into account. Straight translation or the simple adjustment of test items does not express these basic differences. Thus, a ' syllabus-difference ' factor must be allowed for in the interpretation of results. Second, the concept of an ' average mean ' or ' average standard deviation ' between countries seems to entail a supposition that no relative differences exist. In this study these terms seem to have little genuine meaning, and an alternative criterion measure is needed. Such a measure must be capable of rational interpretation. Averages, in this sense, merely obscure the differences they are intended to evaluate. Hotyat's use of profiles is a much more promising technique.

Finally though, it may be possible that certain philosophies of school organisation, such as streaming, may in theory tend to suggest the occurrence of larger standard deviations, it would, with children of secondary school age, be extremely difficult to demonstrate the precedence of this single factor over those described above, or even over straightforward differences in sampling. It is regrettable that the analysis of such excellent raw data took so little account of these and similar problems to the extent that at least one of the conclusions could be thoroughly misleading. D. M. LEE.

GRAY, SUSAN, W. (1963). *The Psychologist in the Schools.* Holt, Rinehart and Winston, Inc., pp. ix+406, $5.50.

Although some parts of this book are likely to be of interest only to the American reader, the main questions dealt with, those relating to the widening functions of the psychologist in the school system, are currently of great concern to L.E A. psychologists in this country, who have become increasingly aware of their role in educational guidance, the prevention of maladjustment and in " the transfer of psychological knowledge, skills and understanding " to others. Our own services can benefit from a knowledge of the ways in which the school psychologist works in the U.S.A., where he has closer links with classroom teachers than is usually the case here.

Unfortunately, the style of the book often irritates and it is over-long, repetitive and full of platitudinous statements. The important sections (which are mainly in the first nine chapters) are obscured by laboured discussions of matters of marginal relevance to the main theme of the book. Nevertheless, there is much practical wisdom, if little that is new, in the author's general treatment of her topic. She exhorts psychologists to try to distinguish between " the urgent and the important " in their professional work and, without casting all school psychologists in the same mould, would like to see them making major contributions to educational progress by initiating and guiding research enquiries in the schools.

The book should encourage educational psychologists, particularly those in training, to give more thought to their professional roles and relationships. There is a useful list of references at the end of each chapter. MAURICE CHAZAN.

JAHODA, MARIE (1963). *The Education of Technologists : An Exploratory Case Study at Brunel College.* London : Tavistock Publications, pp. xiii+226, 30s.

This is an account of the Nuffield Research Project in Technological Education carried out by Dr. Jahoda and her associates at Brunel College of Advanced Technology in the period 1959-1962. The problem consisted in examining the educational process as this was experienced by students working for the Diploma in Technology. The distinctive features of these courses is that students spend a considerable proportion of the four years involved working in industry In Brunel College the ' thin sandwich ' system operates—that is, students alternate six months in college with six months in industry. As may well be imagined, this creates problems of considerable variety ; organisational, educational, personal. Dr. Jahoda discusses most of the problems which confront students, staff and the companies participating in the educational process, in an extremely illuminating way.

The method used in the research project was to interview students on entry, at the end of their first six months in college, and at the end of their first industrial period. The college staff was also interviewed, and a sample of firms. Interview schedules were extremely detailed and sought to throw light on the socio-economic background of students, their previous experience, their professional expectations, detailed impressions of their first industrial period, the integration or otherwise of the college and industrial experiences. A large part of the present report is made up of the analysis of this questionnaire material but some attempt is made throughout to relate it to a wider context of educational principles. Reference is also made to " socio-technical system analysis " as developed at the Tavistock Institute of Human Relations, but this appears as something extraneous to the study.

None of the conclusions from this investigation come as a novel revelation to anyone directly involved in educating technologists—this was not to be expected. It is the function of research in this new field of work to confirm or deny one's subjective impressions and Dr. Jahoda has approached this task carefully and systematically Her findings should be of value in relation to the problems of organising higher education not only in colleges of technology, but also in universities. Within the limits of the method of questionnaire studies, this is a distinctive piece of work, well-designed, carefully analysed and written up in a pleasing style, with a sense of relevance and of the broader issues. JOHN MCLEISH.

KNAPP, B. (1963). *Skill in Sport.* Routledge and Kegan Paul, pp. 204, 30s.

This most readable and worthwhile book has been written by an author with experience both in playing games at the international level and in teaching the theory and practice of physical education to university students. Its aim is to provide teachers and coaches with a basic knowledge of the main factors involved in the acquisition of skill in recreational physical activity. The approach to the subject is psychological, the contents cover in detail the factors affecting the learning of skills, individual differences in acquiring skills, theories of learning and physiological mechanisms. Not only is the coverage broad, but it is fully illustrated by examples taken from a wide and varied number of games and recreational activities. Unfortunately, the material is not organised in an apparently logical sequence and the reviewers feel that many readers will prefer the alternative chapter arrangement suggested by the author.

As this book is primarily intended for the teacher and coach, a selection of examples should have been included to show how learning theories are applied to the organization of teaching material and to the practicalities of teaching. Also, technical terms such as sensory feed-back and the interpretation of sensory cues should have been expanded and illustrated by references to teaching method.

This book does not contain original experimental work by the writer, but its value lies in the commentary and discussion of a prolific number of articles on the acquisition of skill in sport. A number of articles and theories are of doubtful value, but Miss Knapp does not accept the sources uncritically ; for example, the claims made for the transfer effects of training are treated with caution, and the validity of certain tests of phsyical skill is questioned The book's main appeal will be to the student teacher specialising in physical education, not only as a basic textbook on skill learning, but as a useful source book on experimental data.

For a long time there has been a need for a comprehensive work on the learning of physical skills and the author is to be congratulated on her initiative in producing a book which should stimulate both discussion and controlled research in this relatively unexplored field of study. G. R. CROSS and J. C. EVANS.

LYSAUGHT, JEROME P., and WILLIAMS, CLARENCE M. (1963). *A Guide to Programmed Instruction.* New York : Wiley, pp. iii+180, 30s.

There have been so many books written about programmed instruction that it is refreshing to come across one which gets down to tackling the fundamental problem of how to set about and sustain what has been called the ' inspired drudgery' of writing programmes. The success of programmed instruction as a pedagogical technique will depend on the attitudes and skill of those who apply it. Both a favourable attitude and the ability to use the technique fruitfully are more likely to develop amongst teachers if they understand programmed instruction and have tried writing programmes for themselves.

Although rigorously behavioural in approach, this book defines technical terms such as ' reinforcement ' in plain language with abundant examples. In this way, understanding programmed learning is shown to be but a small step from understanding S–R theories on the Thorndikian model. Using the resulting uncluttered reinforcement theory only when absolutely necessary, the authors describe the procedures for generating programmes which they found most effective with neophyte programmers drawn from education and industry.

Three chapters deal with the related preparatory tasks of selecting a suitable fragment of the syllabus, gathering data on the target population, and writing down in behavioural terms the immediate and long-term objectives of the programme. By this stage, the inherent structure of the subject-matter may well be apparent and may help in the selection of an appropriate paradigm from the list which includes linear, branching, conversational chaining, mathematics, and various modifications of these. Each of the paradigms is explained and illustrated : but thereafter the book is concerned solely with linear constructed-response programmes as being more suited to teachers learning the art.

In the chapter on evaluation, potential users of published programmes will find a host of questions likely to embarrass most programmers. But the point is a valid one. A programme is of little use unless the educator knows what it can do. This means that he must be told all about the initial testing of the programme and the assumptions made about the learners' knowledge and abilities.

Yet the intention of this book is not to curb programmers by supplying their critics with ammunition. The authors prefer to welcome newcomers to programming, especially practising teachers, by giving them guidance and encouragement to enable then to set up their own miniature automatic teaching systems which will be self-correcting with a strong bias towards higher efficiency. To all who think that the operation of such a system is likely to be a rewarding activity, this book is highly recommended.

BERNARD DODD.

MICHOTTE, A. (Eng. Trans. by T. R. Miles and Elaine Miles) (1963). *The Perception of Causality.* London : Methuen and Co., xxiii+425 pp., 45s. net.

It is good that we should have this excellent English translation of the great work of Professor Michotte of Louvain. It is, of course, already well known to psychologists who have had many opportunities of hearing Professor Michotte at Congresses as well as reading its French form *La Perception de la Causalité* which was published in 1954. At the same time, however, this work is little known outside strictly psychological circles and its revolutionary finding that there may be direct perception of causation seems to have made no impact on general thinking One may still hear a philosopher saying that no one has ever suggested that a cause may be directly perceived ; from now on there will be no good reason for any well-informed person not knowing that Professor Michotte has suggested this and has directed a massive research supporting the suggestion and demonstrating the conditions under which this perceptual character is experienced.

Michotte rightly emphasises the compelling character of his demonstrations to those who have actually seen them. Under certain restricted spatio-temporal conditions, a moving patch of colour approaching a stationary patch which itself takes on the motion of the first patch after contact is reported to be seen to ' push ' the second object ; ' launching ' it if the first patch stops on contact, ' entraining ' it if the first patch continues to move with the same velocity as the second after impact. The narrowness of the conditions under which this causal experience occurs is no less remarkable than its vividness when it does occur. Lengthening of the interval between impact and resultant movement changes the experience to one in which the second movement is 'triggered' by the first, but is not a continuation of it and finally to an experience in which two distinct movements are seen. Michotte also finds this primary causal experience confined to situations in which there is apparent transfer of movement; he does not find it possible, for example, in such a case as a light going on when a door is opened.

The natural line of explanation of a causal perceptual experience is to attribute it to the after-effect of experiences in which objects have been seen to be set in motion by other objects. Michotte argues srtongly against this view and supports his argument by the results of many ingenious experiments. He agrees, however, in his latest discussion that, in addition to the primary causal experience, there are situations in which we attribute causation as a result of physical experiences, although there are also many physical causal situations in which there is neither the primary nor the secondary type of causal experience. For the nature of the experience of perceiving such secondary and acquired causal dependence, Michotte suggests the term 'perceived dependence' as distinct from 'perceived causality' which is the name given to the phenomenal character of such experiences as the launching or entraining of his basic experiments.

The translation by T. R. and Elaine Miles is clear and adequate. A final commentary by T. R. Miles provides a summary which is a useful guide to the intricacies of Michotte's researches on perceived causality. There are two final essays by T. R. Miles, the first defending Michotte's phenomenology against the suspicion that it offends against neo-behaviouristic orthodoxy by implying the prohibited concepts of 'dualism,' the second discussing the question of how far Michotte's findings contradict what was said about causation by Hume. R. H. THOULESS.

NEWSON, JOHN and ELIZABETH (1963). *Infant Care in an Urban Community.* London: George Allen and Unwin, Ltd., pp. 268, 42s.

This fascinating book records the results of asking 709 Nottingham mothers of year-old children questions about the methods they use in rearing and training their children. The mothers were chosen to form a representative sample and the questions were asked in standard form by health visitors in a half-hour (or longer) interview. The answers are carefully analysed statistically for differences between occupational groups, and the text is considerably enlivened by verbatim replies recorded during some of the interviews.

In an area which has been considered for decades as of fundamental importance for a sound theory of personality development it is surprising that psychologists have been satisfied for so long with the lack of solid information about how, in fact, children *are* being brought up in contrast to the plethora of injunctions as to how they *should* be. As the Newsons suggest, there is need for a psychological ecology in which problems are studied less in laboratories and clinics and to a far greater extent in the natural habitat of the human animal. This is certainly true of early childhood. It is unfortunate that we are so dependent on questionnaire material about infant rearing practices, especially as certain strong biases exist towards giving answers to particular questions in the direction which the subject thinks will be acceptable to the interlocutor. The authors bring this out very clearly in relation, for example, to such things as the use of the 'dummy' as a pacifier, toilet-training, breast-feeding, etc. A large number of pit-falls yawn before the unwary investigator in a study of this kind. The Newsons, being parents themselves, as well as psychologists, are aware of the propensity of parents to lie about their children, and have noted the level of confidence with which their material can be accepted and the nature of the qualifications to be made before generalising from the statistical comparisons.

The class differences which emerge are surprising not so far as their *nature*, but certainly as far as their *amount* is concerned. For example, genital play is checked by 93 per cent. of mothers in the lowest occupational class, whereas 75 per cent. of the highest group ignore it altogether. 80 per cent. of one-year-olds in the lowest group, whose mothers have attempted to train them on the potty, have *never* been successful at twelve months; in the highest occupational group only 36 per cent. belong in this category. 60 per cent. of the lowest working group mothers and fathers practically never go out together during the baby's first year: the corresponding number in the highest group is 25 per cent. Working class mothers breast-feed their children two or three times as rarely as middle-class mothers; they use the dummy as pacifier about twice as often; they smack their one-year-olds rather more often; the fathers participate much less in the process of rearing the child.

The great virtue of this book is that it sticks rather closely to the facts and avoids large generalisations. The authors obviously know the group on which they are reporting intimately. This means that they can often throw a strong light on the interpretation of some remark by placing it in a social and cultural context which belongs to the respondent and not to the recipient of the remark. In short, this is a valuable and sophisticated study; it is strongly to be hoped that preparations are well advanced for further volumes by the same investigators on the later stages of child-handling It would be very acceptable if, supplementing possible questionnaire material, behavioural records could also be made available for the same groups of children and parents. JOHN MCLEISH.

PIAGET, J., INHELDER, B., and SZEMINSKA, A. (Trans. by Lunzer, E. A.) (1960). *The Child's Conception of Geometry*. London: Routledge and Kegan Paul, pp. vii 408, 45s.

This volume must be read in conjunction with Piaget's earlier study of the child's conception of space. The earlier work provides much of the basis from which the authors seek to study the growth of understanding of the more precise concepts involved in measurement and metrical geometry, such as a change in position, conservation and measurement of length, locating points in two or three-dimensional space, and basic notions concerning areas and volumes. The method is similar to that used in previous enquiries, entailing systematic question and answer techniques with individual children of a variety of ages, but continuing up to and beyond the equivalent of the English primary stage. The content of the enquiry is related to the kind of geometry taught in England as against a more intuitive exploration of spatial concepts in the earlier book. It is clearly developmental work giving strong indications of the origins from which the formal study of geometry should arise.

The general lay-out of the results conforms to those found in other Piaget studies as regards a general development through from ego-centric thought and action to the final stage of formal operations. Probably most interesting is the sharp indication from an early stage that the child gains his impressions from the structuring of the three-dimensional situation as a whole, and not from consideration and summation of the separate parts—a statement having far-reaching implications for the early teaching of geometry. Alongside this it is also valuable to note the growth of understanding in conservation of measurement, this being only finally completed at the level of formal operations, where conservation of volume finally becomes crystallised.

While the rigidity of some of the age-ranges given for each stage of thinking might well be queried, the general implications of the whole sequence of results have strong pointers in relation to teaching. Too often the excellent general treatment of spatial concepts in the early years of home and school is followed by a long silence at the junior school level and an abrupt intrusion into formal geometry at the secondary stage Clearly, the main implications of this present work point to a need for continuous fostering and treatment of geometrical ideas so that formal notions are developed as coherent concepts rather than isolated rote laws. In this context, a study of this book is particularly appropriate and valuable for teachers of junior school children. D. M. LEE.

PICKFORD, R. W. (1963). *Pickford Projective Pictures.* London: Tavistock Publications, pp. xi+122, 30s. Picture material 25s. The set 50s.

Those who have tried the various available forms of picture projective tests with young children, will have noticed the gap that exists between the T.A.T. or Symond's Picture Story Test for older children, and the rather childish and off-putting C.A.T. Professor Pickford's new set of pictures may help to fill this gap. The unwillingness of young children to tell lengthy stories, makes it desirable to have a large number of pictures yielding many short stories in which recurrent themes may be discerned.

The Pickford Projective Pictures consist of 120 simple line-drawings which " cover a very wide range of stimulus content," varying from relationships with father figures and authority, and relationships with mother figures including sibling jealousy; to play and conflict between children, and death and ghosts. There are some twenty distinctive themes represented. Professor Pickford suggests that the pictures should be used about six at a time, so that the material will last for about twenty therapeutic sessions. However, for diagnostic use, a larger number of pictures each yielding a short story might well reveal highly significant recurrent themes reflecting some of the child's pre-occupations and worries.

The book accompanying the test material contains some very useful data for responses given to the pictures by fairly large groups of children in Glasgow child guidance clinics. Data of this sort are too rarely available for projective tests.

The book also contains a chapter on the use of the pictures in analytic treatment which must not be allowed to antagonise people who are allergic to Freudian interpretations of innocent comments by children. For example, the comment, " This little boy stole apples from a rough farmer's orchard " is interpreted as a wish to rob the father of his sexuality. This interpretation goes on, for good measure, to interpret the tummy-ache the boy is said to have after eating the apples as a phantasy of incestuous pregnancy, on the parr of the girl telling the story !

With or without interpretations of this kind, the pictures seem to have been used successfully in clinics in Glasgow, and should help to fill a gap in our techniques for persuading children to talk in a revealing way about themselves.

R. HETHERINGTON.

SIMON, BRIAN and JOAN. (Edits.) 1963.
Educational Psychology in the U.S.S.R. London: Routledge and Kegan Paul, pp. xi+283, 40s. net. Translated by Joan Simon.

This book will help the West to understand something of the psychological outlook that informs the whole system of education in the Soviet Union, and give a general picture of the vast scope of Soviet educational psychology. It consists of fourteen papers written by leading Russian workers in this field. Mr. Simon, who made the final selection of papers for translation, says in the Preface that the volume was planned with teachers and educationists particularly in mind.

Part 1 deals with various aspects of the following problems: learning and mental development, education and personality development, mental development and intellectual backwardness, and the role of speech in the regulation of behaviour. In Part 2 there is an interesting account of the changes in outlook regarding the psychology of learning over the period 1900—1960, followed by a more detailed discussion of such topics as the promotion of abstraction and generalization, problem-solving, etc. In the final part of the book there are papers dealing with aspects of the following topics, and others: the psychology of reading, problem-solving in arithmetic, mastery of scientific concepts, mastery and application of concepts in physics, transfer to technical skills.

The reviewer had three main reactions to this book, which he can only mention very briefly. First, the outlook of Soviet and Western educational psychologists is much the same over some main issues. For example, Vygotski's hypothesis that the process of development follows on the process of learning which creates the potential development, is not far from the views of Hebb and perhaps not unrelated to those of Festinger. Again, " a low level of development of verbal-logical thinking results, of course, in little capacity not only for mathematics, but also for a number of other activities "(Krutetski, p. 231) could have been written by British psychometrists. Second, the Russians have paid attention to certain issues that the West has tended to neglect. Elkonin (p 166), quoting Galperin, indicates that the mastering of an action on the plane of speaking aloud aids the growth of mental operations ; as does Leontiev (p. 80). Moreover, Soviet psychologists pay considerable attention to the moving back, as it were, from abstract thinking to practical application (p. 128 and pp. 198 ff.) ; this is a matter of great consequence in polytechnical education. Third, some of the hypotheses are vague, or are proposed on the basis of small numbers of children ; these need verification. For example, exactly how effective are "Pavlovian methods of couterposition " in aiding differentiation in the second signal system so aiding abstraction and generalization (p. 124) ?

This is an important book. It is rich in ideas and cannot be neglected. We are indebted to all those who made the volume available to the English speaking world.

SOLOMON, H. (Edit.) (1961). *Studies in Item Analysis and Prediction.* (Stanford Math. Studies in the Social Sciences, VI) ; U.S.A.: Stanford University Press. London : Oxford University Press, pp. ix+310, 70s.

This is definitely a book for those interested in mathematical theory. It brings together a number of recent contributions to the fields of Item Analysis and Prediction, and develops a coherent picture of the statistical background involved. But the reader must, in general, apply the theory for himself. Apart from a rather brief Editor's introduction, there is little practical help or guidance as to the types of problems which might be solved by the use of the various theoretical aspects discussed.

Whilst such volumes as this could play a considerable role in improving and developing techniques for educational research, they will not achieve their goal until the writers generate a better synthesis of theory and practice. Most practising research workers have neither the background knowledge nor the time to unravel the complexities of advanced statistics before applying it to the problem in hand. Theories of as complex a nature as those propounded in this book need to be illustrated by a variety of very suitable and careful applications. D. M. LEE.

TRAPP, E. P., and HIMELSTEIN, P. (Edits.) (1963). *Readings on the Exceptional Child.* New York: Appleton-Century-Crofts. London: Methuen, pp. 674, 55s.

JENKINS, J. J., and PATERSON, D. G. (Edits.) (1963). *Studies in Individual Differences.* New York: Appleton-Century-Crofts. London: Methuen, pp. 774, 63s.

DE CECCO, J. P. (Edit.) (1963). *Human Learning in the School.* New York: Holt, Reinhart and Winston, pp. 636, $4.75.

Books of readings hold a dubious position among teachers of psychology. Some regard them as undesirable in that they discourage the enquiring student from searching for his material and learning more in his search than that for which he sought. Others regard them as inappropriate unless they are exactly tailored to their own course. Yet others, pointing to the difficulties which arise when large classes of students have their attentions drawn to particular articles, find books of well-chosen readings a great help. Judging by the number of books of readings being produced in America at the moment, there would seem little doubt as to the value placed upon them there.

All these books are good of their kind, they cover different areas of study and illustrate different approaches to their compilation. That by Trapp and Himelstein consists almost entirely of recent, including a few unpublished, papers illustrating such aspects of the study of exceptional children as mental defect, giftedness, sensory handicaps, brain injury and, in a short, rather oddly assorted section, emotional disturbance. This last reflects, no doubt, the paucity of research papers in this field. Three ways of selecting papers were used. Reading lists from 100 American universities were compared to extract a central core of common reading. This proved unhelpful since there was virtually no overlap between them, and many lists were 'grossly outdated and meagre in the experimental area.' A second approach, which was somewhat more successful, was to analyse the current text-books on exceptional children. Most papers were selected on advice of research workers who were approached. A high proportion of the papers will be unknown to most lecturers in this country

Jenkins and Paterson have taken a fairly narrow theme, that of the historical development of the search for intelligence and the understanding of its nature and development. Papers are arranged chronologically and the first, as might have been predicted, is from Galton's *Hereditary Genius*, then follow most of the great names in the history of mental testing: Spearman, Thorndike, Binet, Terman, Thurstone, Kelly, and so on to Vernon, Anastasi, Wechsler, Husen, Cattell and Guilford, covering nearly 100 years. Most of these papers are well known and frequently quoted, together they tell a coherent story of the evolution of the present position with its persistent doubts and its pointers for future research.

The third of these readings, edited by De Cecco, aims at showing the relevance of psychological research and theorising for classroom teaching and learning. Virtually, the whole field of educational psychology is covered in ten sections, containing between three and six papers each. About half the papers are original research reports, the rest are survey articles by well-known workers. Among the topics covered are motivation, programmed learning, problem solving, individual differences and learning in a social setting. There are no papers purely on animal learning and all have a clear application for the classroom. Very often papers under the same heading are written from different viewpoints as, for example, when Burt, Guilford and Nancy Bailey all write under *Intelligence*.

It is difficult to evaluate books of readings, except with reference to specific courses but, for what it is worth, this reviewer would find Jenkins and Paterson a useful adjuct to his book list for students, De Cecco rather less so, and Trapp and Himelstein much less relevant. They are all worth a place in college, department and institute of education libraries. L. B. BIRCH.

STUART, A. (1962). *Basic Ideas of Scientific Sampling*. Charles Griffin and Co., pp. 99, 18s.

The first standard works on sample surveys, such as those of Yates and Deming, appeared between ten and fifteen years ago. Admirable as these books are, they are rather too specialised and technical to be easily absorbed by many psychologists and teachers who need to know the basic principles of sampling. General texts in statistics, on the other hand, usually limit themselves to inference from simple random samples, and this kind of sample is nearly always impracticable in educational research. There has thus been a need for a clear and simple introduction to such methods as cluster sampling, in which the sampling unit may be a class or a school instead of an individual, since relatively few of the research workers who use samples of this kind are completely informed about the principles involved and the methods of analysis required.

Mr. Stuart's monograph should help to bridge this gap. It is succinct, clear, and demands as little mathematical knowledge as any work on a statistical topic can do. There are frequent numerical examples, and these are kept beautifully simple. It is not, however, what the Americans call a 'cook-book,' meaning one that provides rules-of-thumb for the commonest situations without much attention to theory.

The emphasis throughout is theoretical, few formulae are provided, and the main aim is clearly, as suggested by the title, to familiarise readers with the basic logic rather than with the computational complexities of the various sampling processes.

This aim is successfully achieved within the limited compass of the book. There are forty-four sections, averaging two or three pages each, of which the first nineteen deal mainly with general topics such as estimation, the normal distribution and confidence intervals. The last but one goes as far as sampling with probabilities proportional to sizes in cluster and two-stage samples.

Within his self-imposed limits, the author has provided what is probably the clearest available exposition of sampling theory to the non-mathematician, and one that can be studied with benefit by the majority of people doing research in psychology and education. H. J. BUTCHER.

WALL, W. D., SCHONELL, F. J., and OLSON, W. C. (1962). *Failure in School.* Unesco Institute for Education, Hamburg, pp. 158, 11s.

This report is, regrettably, the final one of the series of expert international studies sponsored by Unesco during the last decade. Although the main discussions which culminated in this volume, took place in 1956, the book strives to be up-to-date. Many references are made to recent publications, and the current views of psychologists on school failure and aspects of child development are lucidly if tersely explained. The authors make it clear that they are concerned not only with backward pupils and failure *in* school but with the failure *of* schools and the backwardness of educational systems and practices.

Dr. Wall contributes four valuable chapters, in which he defines the problem, discusses the incidence and consequences of failure, and inquires into the socio-economic, familial and scholastic factors associated with educational retardation. He illuminates his theme with some useful figures, making comparisons between different countries and systems, while not failing to point out the difficulties in the way of obtaining full and reliable information as well as of interpreting the available data. His clear discussions of the significance of discrepancies between ability and attainment (Chapter II) and, in Chapter IV, of the effect of different levels of parental encouragement on attainment—and, indeed, on measured ' g '—are of particular interest.

Professor Olson's single chapter, dealing with aspects of child growth, is excellent. He produces evidence to support his contention that achievement levels must be viewed in the light of a child's total development, and that much that is called ' failure ' is not genuine child failure, but failure administratively defined. The range of individual differences is so wide that the use of class averages as a yardstick is arbitrary and misleading. Like Dr. Tanner in this country, he stresses that chronological age is an unreliable index to a child's educational status or readiness for learning.

Although there is little new in Professor Schonell's two chapters, the first is a guide to child study that teachers and students in training will find extremely helpful, and the second a succinct account of possible ways of organizing remedial measures to help the failing child.

While a report of this kind is bound to lack depth in some places, it will help to clarify thought on school failure and can be recommended to all educationalists, including those who are responsible for the organization of education. There is a comprehensive bibliography which covers publications in eight languages ; the price is reasonable but the binding flimsy. M. CHAZAN.

PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, a later review.

R. F. Tredgold (1963). *Human Relations in Modern Industry.* Gerald Duckworth and Co., pp. 192, 21s.

R. S. Pandey (1963). *Our Adolescents : Their Interests and Education.* Lakshmi Narain Agarwal, Agra, pp. xx+387, 15 rupees.

A. Michotte (1963). *The Perception of Causality.* Methuen and Co., Ltd., pp. xxii+425, 45s.

P. R. Farnsworth (Edit.) (1963). *Annual Review of Psychology, Vol.* 14, 1963. Annual Reviews, Inc., vi+454, $8.50 in U.S.A. ; $9.00 elsewhere.

H Pluckrose (1963). *Picture Making with Juniors.* Oldbourne Press, pp. 141, 18s.

R. Morris (1963). *Success and Failure in Learning to Read.* Oldbourne Press, pp. 176, 17s. 6d.

J. Sandven (Edit.) (1963). *The Role of Educational Research in Social Education*—Papers of the Third International Congress for the University study of Education. Oslo University Press, pp. 302, no price given.

John and Elizabeth Newson (1963). *Infant Care in an Urban Community.* Allen and Unwin, pp 268, 42s.

Virgil E. Herrick (Edit.) (1963). *New Horizons for Research in Handwriting.* The University of Wisconsin Press, pp. xix+276, $5.

Jerome P. Lysaught and Clarence M. Williams (1963). *A Guide to Programmed Instruction.* John Wiley and Sons, pp. ix+180, 30s.

John P. DeCecco (1963). *Human Learning in the School*—Readings in Educational Psychology. New York : Holt, Rinehart and Winston, pp. xix+636, $4.75

Silvan, S. Tomkins and Samuel Messick (Edits) (1963). *Computer Simulation of Personality.* John Wiley and Sons, pp. x+325, 40s.

Calvin W. Taylor and Frank Barron (Edits.) (1963). *Scientific Creativity : Its Recognition and Development.* John Wiley and Sons, pp. xxiv+419, 60s.

D. H Stott (1963, 2nd Edition). *The Social Adjustment of Children.* University of London Press, pp. 63, 25s

Abraham H Maslow (1962) *Towards a Psychology of Being.* D. Van Nostrand Co., Ltd., pp. ix+214, 15s.

Susan W. Gray (1963) *The Psychologist in the Schools.* New York : Holt, Rinehart and Winston, pp ix+406, $5.50.

R. W. Pickford (1963). *Pickford Projective Pictures.* Tavistock Publications, pp. xi+122, test 30s., picture material, 25s., the set 50s.

John H. Flavell (1963). *The Developmental Psychology of Jean Piaget.* Princeton, N. J. : D. Van Nostrand Co., Inc., pp. xvi+472, 68s.

Richard D. Smallwood (1963). *A Decision Structure for Teaching Machines.* Cambridge, Mass. : The M.I T. Press, pp. v+122, 30s.

Winfred F. Hill (1963). *Learning*—A survey of psychological interpretations. San Francisco : Chandler Publishing Co., pp. x+227 (no price given).

Maurice Goldsmith (Edit.) (1963). *Mechanisation in the Classroom.* Souvenir Press, Ltd., pp. 236, 25s.

Kenneth Richmond (1963). *Culture and General Education.* Methuen and Co., pp. 184, 21s

Nicholas Hans (1963). *The Russian Tradition in Education.* Routledge and Kegan Paul, Ltd., pp. vii+196, 30s.

Robert Peers (1963). *Fact and Possibility in English Education.* Routledge and Kegan Paul, Ltd., pp. x+180, 21s.

W. J. Campbell (1963). *Growing up in Karribbee.* Austrialian Council for Educational Research, pp. xvi+147, 30s. Aust.

R. D. Gee (1963). *Teaching Machines and Programmed Learning*—A guide to the literature. A Hertis Publication, pp. 68, 4s.

John P. Wynne (1963). *Theories of Education.* Harper and Row, pp. xiiii+521, 49s.

J. D. Keehn (1963). *Prediction and Control of Behaviour.* Constable, pp. 268, 45s.

Edward Fry (1963). *Teaching Faster Reading*—A Manual. C.U.P., pp. xii+142, 8s. 6d. *Reading Faster*—A Drill Book. C.U.P. pp. xi+67, 3s. 9d.

Lawrence C. Little (1963). *A Bibliography of Doctoral Dissertations on Adults and Audlt Education.* University of Pittsburgh Press, pp. iv+163, $4.50 paper.

Hans Aebli (1963). *Uber die geistige Entwicklung des Kindes.* Stuttgart: Ernst Klett Verlag, pp. 124, 11.80 DM.

Henri Pieron (1963). *Examens et Docimologie.* Presses Universitaires de France, pp. 190, F.9 .

National Association for Mental Health (1963). *Clinical Problems in Children of Primary School Age.* The proceedings of the 19th Child Guidance Inter-Clinic Conference. N.A.M.H., pp. 40, 4s.

Marie Jahoda (1963). *The Education of Technologists.* London: Tavistock Publications, pp. 226, 30s.

H. H. Stern (1963). *Foreign Languages in Primary Education.* Hamburg Unesco Institute for Education, pp. 103.

INDEX TO VOLUME XXXIII, 1963

Articles.

Academic Bias, the Relation of Psychological Test Scores to : L. Hudson 120
Draw-a-man Test, the Reliability and Validity of the Goodenough : M. L K. Pringle and K. T. Pickup 297
Educationally Sub-Normal Boys, Changes in Terman-Merrill I.Qs. of : C. S. Rushton and A. E. Stockwin 132
Ideas about Country and Nationality, the Development of : G. Jahoda :
  I.—The Conceptual Framework 47
  II.—National Symbols and Themes 143
Intra-Family Relationships, a Projective Technique designed for Investigations with Emotionally Disturbed Children : J G. Howells and J. R. Lickorish 286
Juvenile Delinquency, Social Factors and School Influence in : G. A. Dell 312
Leadership in a Dual Cultural Setting · R. K. Muir 253
Learning Ability, the Measurement of : G. W. S. Mackay and P E. Vernon .. 177
Learning Effects of Pictures, the : O. Magne and L. Parknäs 265
Mental Set, an Experiment on the effects of, in the Brook Reaction Test . D. H. Hargreaves, A. W. Heim and K. P. Watts 236
Personality Factors and School Achievement : H. J. Butcher, M Ainsworth and J. E. Nesbitt 276
Personality, Teachers' Description of Six-Year-Olds : A Check on Structure : R. B. Cattell 219
Progressive Matrices (1947), A Further Study of : A. J. Yates 307
Psycho-logic : A Possible Alternative to Piaget's Formulation : G. H McLaughlin 61
Puberty, Early, the Influence of on Test Performance at the Age of Eleven : J. D. Nisbet and R. Illesley 169
Reading, Value of the Oral Response in Beginning · An Experimental Study using Programmed Instruction : J. D. McNeil and E. R. Keislar 162
Reasoning Test for Higher Levels of Intelligence, an Analysis of Valentine's : H. J. Hallworth 41
Remedial Education, a Further Study of the Educational Progress of Children who had received : K Lovell, C. Byrne and B. Richardson 3
Rural Children, are they handicapped by the use of Speeded Test in Selection Procedures ? : C. Anne Moreton and H. J. Butcher 22
Scientific Concepts, the Development of, in Children (2) : W. H King 240
Self Concepts, the, of Trainee Teachers in two Sub-Cultures : A. S. Phillips .. 154
Socioeconomic Bias, the Reduction of, in Intelligence Testing : R S. MacArthur and W. B. Elley 107
Speech Disorder, some Characteristics of Children with Two Types of : T. G. Crookes and M. C. L. Greene 31
Student Performance, Predicting, in a University Department of Education · F. W. Warburton H. J. Butcher and G. M. Forrest 68
Sub-Normal Children, Developmental Patterns of Severely : M. Woodward and Diana J. Stern 10

Critical Notes and Comments :

Social Factors and School Influence in Juvenile Delinquency by G. A. Dell : A. B. Clegg 322
The Effect of Socio-Economic Status on the Scholastic Achievement of Sixth-Grade Children by R. L. Curry . R. R. Dale 84
Specific Dyslexia : D. Russell Davis and Asher Cashdan 80
A Reply : M. D. Vernon 83
On the Relation of Neuroticism and Extraversion to Intelligence and Educational Attainment · H. J. Eysenck 192
Teaching Educational Psychology in Training Colleges : T. H. B. Hollins and J. M. Thyne 187

Research Note :

The Relative Personality Adjustment of Severely Deaf and Partially Deaf Children : L. R. Bowyer, A. Marshall and K. Weddell 85

Critical Notices :

Intelligence and Experience by J. McVie Hunt : P. E. Vernon 193
The Appreciation of Beauty by C. W. Valentine : C. Burt 194

Authors:

Butcher, H. M., Ainsworth, M., and Nesbitt, J. E.: Personality Factors and School Achievement 276
Cattell, R. B.: Teachers' Personality Description of Six-Year-Olds: A Check on Structure 219
Crookes, T. G., and Greene, M. C. L.: Some Characteristics of Children with Two Types of Speech Disorder 31
Dell, G. A.: Social Factors and School Influence in Juvenile Delinquency 312
Hallworth, H. M.: An Analysis of C. W. Valentine's Reasoning Test for Higher Levels of Intelligence 41
Hargreaves, D. H., Heim, A. W., and Watts, K. P.: An Experiment on the Effects of Mental Set in the Brook Reaction Test 236
Howells, J. G., and Lickorish, J. R.: A Projective Technique for Investigating Intra-family Relationships designed for use with Emotionally Disturbed Children 286
Hudson, L.: The Relation of Psychological Test scores to Academic Bias 120
Jahoda, G.: The Development of Children's Ideas about Country and Nationality:
  I.—The Conceptual Framework 47
  II.—National Symbols and Themes 143
King, W. H.: The Development of Scientific Concepts in Children (2) 240
Lovell, K., Byrne, C., and Richardson, B.: A Further Study of the Educational Progress of Children who had received Remedial Education 3
MacArthur, R. S., and Elley, W. B.: The Reduction of Socio-economic Bias on Intelligence Testing 107
Mackay, G. W. S., and Vernon, P. E.: The Measurement of Learning Ability 177
McLaughlin, G. H.: Psycho-Logic: A Possible Alternative to Piaget's Formulation 61
McNeil, J. D., and Keislar, E. R.: Value of the Oral Response in Beginning Reading: An Experimental Study using Programmed Instruction 162
Magne, O., and Parknäs, L.: The Learning Effects of Pictures 265
Moreton, C. A., and Butcher, H. J.: Are Rural Children Handicapped by the Use of Speeded Tests in Selection Procedures? 22
Muir, R. K.: Leadership in a Dual Cultural Setting 253
Nisbet, J. D., and Illesley, R.: The Influence of Early Puberty on Test Performance at the Age of Eleven 169
Phillips, A. S.: The Self Concepts of Trainee Teachers in Two Sub-Cultures 154
Pringle, M. L. K., and Pickup, K. T.: The Reliability and Validity of the Goodenough Draw-a-Man Test 297
Rushton, C. S., and Stockwin, A. E.: Changes in Terman-Merrill I.Qs. of Educationally Sub-Normal Boys 132
Warburton, F. W., Butcher, H. J., and Forrest, G. M.: Predicting Student Performance in a University Department of Education 68
Woodward, M., and Stern, D. J.: Developmental Patterns of Severely Sub-Normal Children 10
Yates, A. J.: A Further Study of Progressive Matrices (1947) 307

NOTES TO CONTRIBUTORS

1.—The Editor is glad to consider articles in the fields of educational and child psychology, whether written from the theoretical, clinical, experimental or psychometric standpoints. It is, however, desirable that statistical matter (formulae, tables, discussions of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers Such concepts as the mean, percentiles, standard deviation, correlation, chi-squared, standard error, critical ratio, significance and reliability, and general or group factors, can be employed in the text. But more complex concepts and methods should be explained or confined to footnotes or an appendix, or published elsewhere.

Articles which have been, or are to be, published elsewhere cannot be accepted.

2.—*Length.* The usual range of length is from 2,500 to 4,500 words, and only very exceptionally will articles of more than 6,000 words be published. Short research reports of 500—1,000 words, such as outlines of theses, are welcomed. These should normally be submitted through the Head of the Department in which the research was carried out.

3.—*Arrangement.* Articles should be *preceded* by a summary or abstract of between 100 and 300 words, embodying the main conclusions. This should be typed directly after the title and author's name and place of work. A contents analysis, or list of headings, need *not* be included, but successive sections of the article should receive numbered headings, e.g. ·

I. INTRODUCTION.
II. OUTLINE OF PREVIOUS RESEARCH.
III. PLAN OF THE INVESTIGATION.
  (i) *Subjects and materials.*
  (ii) *Experimental Procedure.*
IV. STATEMENT OF RESULTS.
V. DISCUSSION AND CONCLUSIONS.
VI. REFERENCES.

4 —*Typing.* All papers must be typed (double or 1½-spaced) on one side of quarto or foolscap paper. There should be margins of at least one inch all round Authors should consult articles in current numbers of the *Journal*, and copy their conventions in respect of capitalising the title and main section headings, italicising author's place of work and sub-section headings, indentation of paragraphs, the form of references and tables, etc.

5.—*Footnotes* should be few in number, and should be indicated in the text by using the signs *, †, ‡, §.

6.—*References* should normally be indicated by date following the author's name in the text, e.g., Spearman (1904), or (Spearman, 1904, p. 160), or—where there are two or more publications in the same year—Burt (1939a). However, it is occasionally more convenient to number the references throughout, and to type numbers in the text above the line, e.g , Spearman.[4] When there are two or three references only, these may be placed in footnotes in the text, instead of in a separate section at the end.

The list of references at the end should always be in alphabetical order of author's names, and—for any one author—by order of date The following standard form of references should be observed :

DEARBORN, W., F., and ROTHNEY, J. W. M. (1941). *Predicting the Child's Development* Cambridge, Mass. : Sci.-Art.

(Note that Ltd., etc., after publisher's names, also the numbers of pages in the book, can be omitted.)
BURT, C. (1939a). The relations of educational abilities. *Brit. J. Educ Psychol ,* 9, 55-71.

(Note the absence of inverted commas and of capital initials in the titles of articles. The number of the volume is in arabic, not roman numerals. The part of the journal is not given, but page numbers are included in full).

The World List of abbreviations (as given in the December issues of *Psychological Abstracts*) is used for journals.

7.—*Tables and figures* are costly to print, and should be used only when they save space and/or are essential for effective presentation. The same material should not be presented in a table *and* a graph. Each table or figure should be numbered (in arabic figures), on a separate sheet, and its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy, unruled paper, the lettering on graphs being in pencil. Lines in tables should be ruled in pencil. Each table or graph should have a short title, sufficient for understanding without searching in the text.

8.—*Corrections.* The Editor cannot consider for publication articles which are seriously deficient in presentation. The typescript submitted (after preliminary consultation, if desired) should represent the final form in which the author wishes the paper to appear. Since changes in proof are costly, out of all proportion to the original setting, authors must be charged for any changes, insertions or deletions other than printer's errors).

9 —*Reprints.* Authors receive fifty copies of their papers free Extra copies, in multiples of fifty may be had at cost price if the order is given when the proof is returned.

E. & S. LIVINGSTONE, LTD.

THE DEVELOPMENT OF THE INFANT AND YOUNG CHILD. Normal and Abnormal.

Second Edition. By R. S. Illingworth, M.D., F.R.C.P., D.P.H., D.C.H.

370 pages. 126 illustrations. 30s.

"A splendid simplification and condensation of the original studies of Gesell and others adapted by the author for day-to-day clinical use and widely modified by his own large experience. The book is entertaining and it is practical; it is very well written; it should certainly be read by all who are engaged in the study of young children."

Archives of Disease in Childhood.

THE NATURAL DEVELOPMENT OF THE CHILD

Fourth Edition. By Agatha H. Bowley, Ph.D., F.B Psy.S.

224 pages. 84 illustrations. 10s 6d.

"This book, written primarily as a guide for teachers and students, has come to be recognised as a standard work for all those concerned with child care."

Woman Health Officer.

TEVIOT PLACE, EDINBURGH

THE
BIRMINGHAM
PRINTERS LTD
PRINTERS & BOOKBINDERS

High Class
Technical
Publications

42-44 HILL STREET
BIRMINGHAM 5
TELEPHONE: MIDLAND 0196-7

OCCUPATIONAL PSYCHOLOGY

Editor: ALEC RODGER

JULY, 1963 Volume 37, No. 3

CONTENTS

A Study of Frustration and Aggression Among Factory and Office Workers. By DENIS PYM.

A Preliminary Enquiry into Engineering Drawing Comprehension. By J. SPENCER.

Age of Manual Workers in Relation to Conditions and Demands of Work. By MICHAEL S. FEATHERSTONE and CATHERINE M. CUNNINGHAM.

Traditional Values and Work Efficiency in a West African Mine Labour Force. By J. L. M. DAWSON.

Prescription and Perception of the Supervisory Role. By BRIAN R. KAY.

The Consistency of Meaning of Everyday Quantitative Expressions. By D. SHEPPARD.

Other Books Received.

Annual Subscription 40 shillings.

National Institute of Industrial Psychology

14, Welbeck Street, London, W.1.

THE BRITISH JOURNAL OF MEDICAL PSYCHOLOGY

Edited by Dr THOMAS FREEMAN

(*Contents of Vol. XXXVI, Part* 3, *September*, 1963)

W. H. GILLESPIE. Some Regressive Phenomena in Old Age.
J. E. WEBLIN. Psychogenesis in Asthma.
IRVING L. JANIS. Group Identification under Conditions of External Danger.
HELMUT THOMAE. Some Psycho-Analytic Observations on Anorexia Nervosa.
LINTON GRIMSHAW. Anorexia Nervosa—Progress of a Case
JAMES CHAPMAN and ANDREW MCGHIE. An Approach to the Psychotherapy of Cognitive Dysfunction in Schizophrenia.
ROBIN HIGGINS. The Perception of the Body Surface.
GWEN DOUGLAS. Puerperal Depression and Excessive Compliance with the Mother.
WILLIAM R. PIERSON. Fatigue, Work Decrement and Endurance in a Simple, Repetitive Task.
JANE MORRIS. An Analysis of Schizophrenics' Interviews.
REVIEWS.
CORRESPONDENCE.

Subscription price 75s. net per volume of 4 parts. Single parts 25s. net.

Orders may be placed with any bookseller, or sent direct to :
CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200, Euston Road, London, N.W.1

COMMUNIST EDUCATION

Edited by E. J. King

Recent developments and current trends in communist education are traced in this survey made by specialists who have studied and visited the U.S.S.R. and other communist countries. **25s.**

CULTURE AND GENERAL EDUCATION

Kenneth Richmond

The results of a 'test of culture' devised by the author are examined and discussed. They shed light on present standards of education and on the 'great divide' between the scientist and the arts man. **21s.**

METHUEN *& Co. Ltd. · 36 Essex Street · Strand · London W.C.2*

Sixth Edition now available

EDUCATIONAL PSYCHOLOGY AND CHILDREN

K. Lovell, M.A., B.Sc., Ph.D.

This is a comprehensive and up-to-date textbook of Educational Psychology. The author has met the examination requirements of two-year and post-graduate students training to be teachers, and of those commencing a more advanced course in Education and taking Educational Psychology as a subject. At the same time the book will be of great value to practising teachers, social workers, members of the medical profession, and to all concerned with the upbringing of children. This new edition has been thoroughly revised and all references to literature have been brought up to date.

Sixth Edition, 18*s net*

UNIVERSITY OF LONDON PRESS LTD
WARWICK SQUARE LONDON EC4

Language, Thought and Personality

in Infancy and Childhood

M. M. LEWIS, M.A., Ph.D.

The intention of this book is to give a systematic account of the growth of language in relation to general development in infancy and childhood. Although over the past thirty years much work has been done on different aspects of language and of its place in human behaviour, little or nothing has been attempted to give an ordered account of the progressive development of language in relation to general growth from birth throughout childhood.

Ready November. *E* 18*s. net.*

HARRAP *182 High Holborn, London, W.C.1*

WILEY
Interscience

Handbook of Mathematical Psychology

Complete in 3 volumes, Volume I
edited by R. D. Luce, R. R. Bush, *University of Pennsylvania*, and E Galanter, *University of Washington.*

The first attempt on such a large scale to describe systemically the major areas of psychology to which mathematical methods have been seriously applied. Most of the chapters are primarily concerned with traditional psychological topics treated from a mathematical point of view. Some, however, are primarily methodological. The first volume includes measurement and psycho-physics and, also, two chapters of methodology. The reader should have a background in calculus, and probability theory and applications. 504 *pages*, 80*s*.

Readings in Mathematical Psychology

Complete in 2 volumes, Volume I.
edited by R D. Luce, R. R. Bush, *University of Pennsylvania*, and E. Galanter, *University of Washington.*

This book of readings is designed to be used in conjunction with the *HANDBOOK OF MATHEMATICAL PSYCHOLOGY*, by the same editors. It assembles into two convenient volumes articles in this field from various journals, not readily accessible to the research worker or student in foreign countries 545 *pages*, 68*s*.

Scientific Creativity:

Its Recognition and Development

edited by C. W. Taylor, *Chairman of the Utah Conferences and Professor of Psychology, University of Utah*, and F. Barron, *Research Psychologist, University of California, and Director, The Berkeley Research Institute.*

Contains the Selected Papers of the first three University of Utah Conferences : "The Identification of Creative Scientific Talent." The relation of comprehensive research findings to education is discussed. In trying to understand the nature of creative talent and how to identify it, intellectual, motivational, biographical, sociological and other personality characteristics are treated. 443 *pages*, 60*s*.

Six Cultures:

Studies of Child Rearing

edited by Beatrice B. Whiting, *Laboratory of Human Development, Harvard University*, with the assistance of I. L. Child, *Yale University*, W. W. Lambert, *Cornell University*, and J. W. M. Whiting, *Harvard University.*

This volume is the first publication reporting world-wide field studies undertaken in 1954 by a group of social scientists from Harvard, Yale, and Cornell Universities. In its broadest conception, the research was aimed at exploring cross-culturally the relation between different patterns of child rearing, socialization and the subsequent differences in personality. 1023 *pages*, 95*s*.

John Wiley · Interscience
Glen House · Stag Place · London SW1

SOCIOLOGICAL REVUE MONOGRAPH No. 7

Sociological Studies in British University Education

October, 1963

Edited by Dr. PAUL HALMOS

CONTRIBUTIONS

Introduction	Sir Eric Ashby
University-State Relations Re-Examined	Dr. R. O. Berdahl
The State and the Universities	Professor J. Blondel
The Pool of Ability	Professor P. E. Vernon
The Too Few Chosen, and the Many that Could be Called	Mr. W. D. Furneaux
Student Selection—Implications Derived from Two Student Selection Enquiries	Dr. Hilde Himmelweit
University Student Selection in Relation to Subsequent Academic Performance; a Critical Appraisal of the British Evidence	Professor R. K. Kelsall
Some Factors Related to the Academic Performance of British University Students	Mr. J. G. H. Newfield
Reflections on the Influence of Social Class on Student Performance at the University	Mr. R. R. Dale
The Influence of Emotional Factors on Achievement in University Education	Dr. Nicolas Malleson
Oxbridge and Redbrick: the Debt and the Interest	Professor W. R. Niblett
The Universities' Role in Higher Education	Professor Boris Ford
Some Myths of University Expansion	Dr. A. N. Little
The Sociology of University Education	Professor T. S. Simey

Copies can be ordered from :

The Secretary to the MONOGRAPH Editor, Department of Psychology, University of Keele, Keele, Staffs., ENGLAND

Price 32s. plus 9d. postage. U.S.: $5.00 plus 15c postage

OTHER PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology.
The British Journal of Psychology (Monograph Supplements).
The British Journal of Medical Psychology.
The British Journal of Statistical Psychology
The British Journal of Social and Clinical Psychology.

The British Journal of Psychology and *The British Journal of Medical Psychology* appear in quarterly parts in volumes of about 350 pages, subscription price of each, payable in advance, 60s. per volume (post free). Subscriptions may be sent to any Bookseller or to the Cambridge University Press, Bentley House, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $10.00. Enquiries should be addressed to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22.

The British Journal of Statistical Psychology is issued twice a year, in May and November, in volumes of about 184 pages. The subscription rate payable in advance is 30s. 6d. per volume (post free). The subscription price in the U.S.A. is $4.50 (post free). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, E.C.4., or to any Bookseller.

The British Journal of Social and Clinical Psychology is issued three times a year in February, June and November, in volumes of approximately 240 pages. The subscription price of each, payable in advance, is 50s. (post free). Subscriptions may be sent to any Bookseller, or to the Cambridge University Press, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $8.50. Enquiries should be sent to the Cambridge University Press American Branch, 32, East 57th Street, New York 22. Separate parts will be sold at 20s.: $3.50 in the United States.

A series of *Monograph Supplements* is issued in connection with the *British Journal of Psychology*, but is not included in the subscription for that journal. Orders may be sent to any Bookseller or to the Cambridge University Press, 200, Euston Road, London, N.W.1., from whom details of these publications may be obtained.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to .
Dr. BORIS SEMEONOFF, Department of Psychology, The University, Edinburgh.

For the *British Journal of Medical Psychology*, should be sent to Dr. THOMAS FREEMAN, Lansdowne Clinic, 4, Royal Crescent, Kelvingrove Glasgow, C.3.

For the *British Journal of Educational Psychology* should be sent to Mr. L. B. BIRCH, Institute of Education, The University, Sheffield, 10.

For the *British Journal of Statistical Psychology*, should be sent to DR. R. J. AUDLEY, Department of Psychology, University College, Gower Street, London, W C.1.

For the *British Journal of Social and Clinical Psychology* should be sent to Mr MICHAEL ARGYLE (Social Psychology Editor), Institute of Experimental Psychology, 1, South Parks Road, Oxford ;

or to :

Dr. JACK TIZARD (Clinical Psychology Editor), Maudsley Hospital, Denmark Hill, London S.E.5.

For the *Monograph Supplements*, to Professor L. S. HEARNSHAW, Department of Psychology, The University, 7, Abercromby Square, Liverpool, 7, whose decision regarding the publication of all *Monograph Supplements* will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1st September of the preceding year.

OFF-PRINTS

Contributors receive 50 free copies of their article and may buy more in units of 50.

Members of the British Psychological Society receive all the Society's Journals at reduced rates.

Information on membership and further activities may be obtained from the Secretary, Tavistock House South, Tavistock Square, London, W.C.1.

PUBLISHED BY THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May and distributed free to members of the Association. To non-members it costs, 5s. per issue, post free (subscription 14s. per annum).

Members of the A.T.C.D.E. receive the *British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151, Gower Street, London, W.C.1.

CONTENTS

		PAGE
R. B. CATTELL:	Teachers' Personality Description of Six-Year-Olds: A Check on Structure	219
D. H. HARGREAVES, A. W. HEIM and K. P. WATTS:	An Experiment on the Effects of Mental Set in the Brook Reaction Test	236
W. H. KING:	The Development of Scientific Concepts in Children	240
R. K. MUIR:	Leadership in a Dual Cultural Setting	253
O. MAGNE and L. PARKNÄS:	The Learning Effects of Pictures	265
H. J. BUTCHER, M. AINSWORTH and J. E. NESBITT:	Personality Factors and School Achievement	276
J. G. HOWELLS and J. T. LICKORISH:	A Projective Technique for Investigating Intra-Family Relationships Designed for use with Emotionally Disturbed Children	286
M. L. K. PRINGLE and K. T. PICKUP:	The Reliability and Validity of the Goodenough Draw-A-Man Test	297
A. J. YATES:	A Further Study of Progressive Matrices (1947)	307
G. A. DELL:	Social Factors and School Influence in Juvenile Deli[illegible]ency	312
A. B. CLEGG:	Comments on Dell's Article.	322
SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES		323
BOOK REVIEWS		328
PUBLICATIONS RECEIVED		339
INDEX TO VOLUME XXXIII		341
NOTES TO CONTRIBUTORS		343

11 DEC 196